Solutions to Linear Algebra, Fourth Edition, Stephen H. Friedberg, Arnold J. Insel, Lawrence E. Spence

Jephian Lin, Shia Su, Zazastone Lai

July 27, 2011

Prologue

This is Solution to Linear Algebra written by Friedberg, Insel, and Spence. And this file is generated during the Linear Algebra courses in Fall 2010 and Spring 2011. I was a TA in these courses. Although this file will be uploaded to the course website for students, the main purpose to write the solution is to do some exercises and find some ideas about my master thesis, which is related to some topic in graph theory called the *minimum rank problem*.

Here are some important things for students and other users. The first is that there must be several typoes and errors in this file. I would be very glad if someone send me an email, jlch3554@hotmail.com, and give me some comments or corrections. Second, for students, the answers here could not be the answer on any of your answer sheet of any test. The reason is the answers here are simplistic and have some error sometimes. So it will not be a good excuse that your answers is as same as answers here when your scores flied away and leave you alone.

The file is made by MikTex and Notepad++ while the graphs in this file is drawn by IPE. Some answers is mostly computed by wxMaxima and little computed by WolframAlpha. The English vocabulary is taught by Google Dictionary. I appreciate those persons who ever gave me a hand including people related to those mentioned softwares, those persons who buttressed me, and of course those instructors who ever taught me.

Thanks.

<div align="right">

-Jephian Lin
Department of Mathematrics, National Taiwan University
2011, 5/1

</div>

A successful and happy life requires life long hard working.

<div align="right">

Prof. Peter Shiue

</div>

Version Info

- 2011, 7/27—First release with GNU Free Documentation License.

Contents

Chapter 1

Vector Spaces

1.1 Introduction

1. (a) No. $\frac{3}{6} \neq \frac{1}{4}$

   (b) Yes. $-3(-3, 1, 7) = (9, -3, -21)$

   (c) No.

   (d) No.

2. Here t is in $\mathbb{F}$.

   (a) $(3, -2, 4) + t(-8, 9, -3)$

   (b) $(2, 4, 0) + t(-5, -10, 0)$

   (c) $(3, 7, 2) + t(0, 0, -10)$

   (d) $(-2, -1, 5) + t(5, 10, 2)$

3. Here s and t are in $\mathbb{F}$.

   (a) $(2, -5, -1) + s(-2, 9, 7) + t(-5, 12, 2)$

   (b) $(3, -6, 7) + s(-5, 6, -11) + t(2, -3, -9)$

   (c) $(-8, 2, 0) + s(9, 1, 0) + t(14, 3, 0)$

   (d) $(1, 1, 1) + s(4, 4, 4) + t(-7, 3, 1)$

4. Additive identity, 0, should be the zero vector, $(0, 0, \ldots, 0)$ in $\mathbb{R}^n$.

5. Since $x = (a_1, a_2) - (0, 0) = (a_1, a_2)$, we have $tx = (ta_1, ta_2)$. Hence the head of that vector will be $(0, 0) + (ta_1, ta_2) = (ta_1, ta_2)$.

6. The vector that emanates from (a, b) and terminates at the midpoint should be $\frac{1}{2}(c - a, d - b)$. So the coordinate of the midpoint will be $(a, b) + \frac{1}{2}(c - a, d - b) = ((a + c)/2, (b + d)/2)$.

7. Let the four vertices of the parallelogram be A, B, C, D counterclockwise. Say $x = \vec{AB}$ and $y = \vec{AD}$. Then the line joining points B and D should be $x + s(y - x)$, where s is in $\mathbb{F}$. And the line joining points A and C should be $t(x + y)$, where t is in $\mathbb{F}$. To find the intersection of the two lines we should solve s and t such that $x + s(y - x) = t(x + y)$. Hence we have $(1 - s - t)x = (t - s)y$. But since x and y can not be parallel, we have $1 - s - t = 0$ and $t - s = 0$. So $s = t = \frac{1}{2}$ and the midpoint would be the head of the vector $\frac{1}{2}(x + y)$ emanating from A and by the previous exercise we know it's the midpoint of segment AC or segment BD.

1.2 Vector Spaces

1. (a) Yes. It's condition (VS 3).

   (b) No. If x, y are both zero vectors. Then by condition (VS 3) $x = x + y = y$.

   (c) No. Let e be the zero vector. We have $1e = 2e$.

   (d) No. It will be false when $a = 0$.

   (e) Yes.

   (f) No. It has m rows and n columns.

   (g) No.

   (h) No. For example, we have that $x + (-x) = 0$.

   (i) Yes.

   (j) Yes.

   (k) Yes. That's the definition.

2. It's the 3×4 matrix with all entries $=0$.

3. $M_{13} = 3$, $M_{21} = 4$, $M_{22} = 5$.

4. (a) $\begin{pmatrix} 6 & 3 & 2 \\ -4 & 3 & 9 \end{pmatrix}$.

   (b) $\begin{pmatrix} 1 & -1 \\ 3 & -5 \\ 3 & 8 \end{pmatrix}$.

   (c) $\begin{pmatrix} 8 & 20 & -12 \\ 4 & 0 & 28 \end{pmatrix}$.

   (d) $\begin{pmatrix} 30 & -20 \\ -15 & 10 \\ -5 & -40 \end{pmatrix}$.

   (e) $2x^4 + x^3 + 2x^2 - 2x + 10$.

   (f) $-x^3 + 7x^2 + 4$.

(g) $10x^7 - 30x^4 + 40x^2 - 15x$.

(h) $3x^5 - 6x^3 + 12x + 6$.

5. $\begin{pmatrix} 8 & 3 & 1 \\ 3 & 0 & 0 \\ 3 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 9 & 1 & 4 \\ 3 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 17 & 4 & 5 \\ 6 & 0 & 0 \\ 4 & 1 & 0 \end{pmatrix}$.

6. $M = \begin{pmatrix} 4 & 2 & 1 & 3 \\ 5 & 1 & 1 & 4 \\ 3 & 1 & 2 & 6 \end{pmatrix}$. Since all the entries has been doubled, we have $2M$ can describe the inventory in June. Next, the matrix $2M - A$ can describe the list of sold items. And the number of total sold items is the sum of all entries of $2M - A$. It equals 24.

7. It's enough to check $f(0) + g(0) = 2 = h(0)$ and $f(1) + g(1) = 6 = h(1)$.

8. By (VS 7) and (VS 8), we have $(a + b)(x + y) = a(x + y) + b(x + y) = ax + ay + bx + by$.

9. For two zero vectors 0_0 and 0_1, by Thm 1.1 we have that $0_0 + x = x = 0_1 + x$ implies $0_0 = 0_1$, where x is an arbitrary vector. If for vector x we have two inverse vectors y_0 and y_1. Then we have that $x + y_0 = 0 = x + y_1$ implies $y_0 = y_1$. Finally we have $0a + 1a = (0 + 1)a = 1a = 0 + 1a$ and so $0a = 0$.

10. We have sum of two differentiable real-valued functions or product of scalar and one differentiable real-valued function are again that kind of function. And the function $f = 0$ would be the 0 in a vector space. Of course, here the field should be the real numbers.

11. All condition is easy to check because there is only one element.

12. We have $f(-t) + g(-t) = f(t) + g(t)$ and $cf(-t) = cf(t)$ if f and g are both even function. Futhermore, $f = 0$ is the zero vector. And the field here should be the real numbers.

13. No. If it's a vector space, we have $0(a_1, a_2) = (0, a_2)$ be the zero vector. But since a_2 is arbitrary, this is a contradiction to the uniqueness of zero vector.

14. Yes. All the condition are preserved when the field is the real numbers.

15. No. Because a real-valued vector scalar multiply with a complex number will not always be a real-valued vector.

16. Yes. All the condition are preserved when the field is the rational numbers.

17. No. Since $0(a_1, a_2) = (a_1, 0)$ is the zero vector but this will make the zero vector not be unique, it cannot be a vector space.

18. No. We have $((a_1, a_2) + (b_1, b_2)) + (c_1, c_2) = (a_1 + 2b_1 + 2c_1, a_2 + 3b_2 + 3c_2)$ but $(a_1, a_2) + ((b_1, b_2) + (c_1, c_2)) = (a_1 + 2b_1 + 4c_1, a_2 + 3b_2 + 9c_2)$.

19. No. Because $(c+d)(a_1, a_2) = ((c+d)a_1, \frac{a_2}{c+d})$ may not equal to $c(a_1, a_2) + d(a_1, a_2) = (ca_1 + dc_1, \frac{a_2}{c} + \frac{a_2}{d})$.

20. A sequence can just be seen as a vector with countable-infinite dimensions. Or we can just check all the condition carefully.

21. Let 0_V and 0_W be the zero vector in V and W respectly. Then we have $(0_V, 0_W)$ will be the zero vector in Z. The other condition could also be checked carefully. This space is called the direct product of V and W.

22. Since each entries could be 1 or 0 and there are $m \times n$ entries, there are $2^{m \times n}$ vectors in that space.

1.3   Subspaces

1. (a) No. This should make sure that the field and the operations of V and W are the same. Otherwise for example, $V = \mathbb{R}$ and $W = \mathbb{Q}$ respectly. Then W is a vector space over $\mathbb{Q}$ but not a space over $\mathbb{R}$ and so not a subspace of V.

   (b) No. We should have that any subspace contains 0.

   (c) Yes. We can choose $W = 0$.

   (d) No. Let $V = \mathbb{R}$, $E_0 = \{0\}$ and $E_1 = \{1\}$. Then we have $E_0 \cap E_1 = \varnothing$ is not a subspace.

   (e) Yes. Only entries on diagonal could be nonzero.

   (f) No. It's the summation of that.

   (g) No. But it's called isomorphism. That is, they are the same in view of structure.

2. (a) $\begin{pmatrix} -4 & 5 \\ 2 & -1 \end{pmatrix}$ with tr= -5.

   (b) $\begin{pmatrix} 0 & 3 \\ 8 & 4 \\ -6 & 7 \end{pmatrix}$.

   (c) $\begin{pmatrix} -3 & 0 & 6 \\ 9 & -2 & 1 \end{pmatrix}$.

   (d) $\begin{pmatrix} 10 & 2 & -5 \\ 0 & -4 & 7 \\ -8 & 3 & 6 \end{pmatrix}$ with tr= 12.

   (e) $\begin{pmatrix} 1 \\ -1 \\ 3 \\ 5 \end{pmatrix}$.

9

(f) $\begin{pmatrix} -2 & 7 \\ 5 & 0 \\ 1 & 1 \\ 4 & -6 \end{pmatrix}$.

(g) $\begin{pmatrix} 5 & 6 & 7 \end{pmatrix}$.

(h) $\begin{pmatrix} -4 & 0 & 6 \\ 0 & 1 & -3 \\ 6 & -3 & 5 \end{pmatrix}$ with tr= 2.

3. Let $M = aA + bB$ and $N = aA^t + bB^t$. Then we have $M_{ij} = aA_{ij} + bB_{ij} = N_{ji}$ and so $M^t = N$.

4. We have $A^t_{ij} = A_{ji}$ and so $A^t_{ji} = A_{ij}$.

5. By the previous exercises we have $(A + A^t)^t = A^t + (A^t)^t = A^t + A$ and so it's symmetric.

6. We have that $\mathrm{tr}(aA + bB) = \sum_{i=1}^n aA_{ii} + bB_{ii} = a\sum_{i=1}^n A_{ii} + b\sum_{i=1}^n B_{ii} = a\mathrm{tr}(A) + b\mathrm{tr}(B)$.

7. If A is a diagonal matrix, we have $A_{ij} = 0 = A_{ji}$ when $i \neq j$.

8. Just check whether it's closed under addition and scalar multiplication and whether it contains 0. And here s and t are in $\mathbb{R}$.

   (a) Yes. It's a line $t(3, 1, -1)$.

   (b) No. It contains no $(0, 0, 0)$.

   (c) Yes. It's a plane with normal vector $(2, -7, 1)$.

   (d) Yes. It's a plane with normal vector $(1, -4, -1)$.

   (e) No. It contains no $(0, 0, 0)$.

   (f) No. We have both $(\sqrt{3}, \sqrt{5}, 0)$ and $(0, \sqrt{6}, \sqrt{3})$ art elements of W_6 but their sum $(\sqrt{3}, \sqrt{5} + \sqrt{6}, \sqrt{3})$ is not an element of W_6.

9. We have $W_1 \cap W_3 = \{0\}$, $W_1 \cap W_4 = W_1$, and $W_3 \cap W_4$ is a line $t(11, 3, -1)$.

10. We have W_1 is a subspace since it's a plane with normal vector $(1, 1, \ldots, 1)$. But this should be checked carefully. And since $0 \notin W_2$, W_2 is not a subspace.

11. No in general but Yes when $n = 1$. Since W is not closed under addition. For example, when $n = 2$, $(x^2 + x) + (-x^2) = x$ is not in W.

12. Directly check that sum of two upper triangular matrix and product of one scalar and one upper triangular matrix are again uppe triangular matrices. And of course zero matrix is upper triangular.

13. It's closed under addition since $(f + g)(s_0) = 0 + 0 = 0$. It's closed under scalar multiplication since $cf(s_0) = c0 = 0$. And zero function is in the set.

14. It's closed under addition since the number of nonzero points of $f + g$ is less than the number of union of nonzero points of f and g. It's closed under scalar multiplication since the number of nonzero points of cf equals to the number of f. And zero function is in the set.

15. Yes. Since sum of two differentiable functions and product of one scalar and one differentiable function are again differentiable. The zero function is differentiable.

16. If $f^{(n)}$ and $g^{(n)}$ are the nth derivative of f and g. Then $f^{(n)} + g^{(n)}$ will be the nth derivative of $f + g$. And it will continuous if both $f^{(n)}$ and $g^{(n)}$ are continuous. Similarly $cf^{(n)}$ is the nth derivative of cf and it will be continuous. This space has zero function as the zero vector.

17. There are only one condition different from that in Theorem 1.3. If W is a subspace, then $0 \in W$ implies $W \neq \varnothing$. If W is a subset satisfying the conditions of this question, then we can pick $x \in W$ since it't not empty and the other condition assure $0x = 0$ will be a element of W.

18. We may compare the conditions here with the conditions in Theorem 1.3. First let W be a subspace. We have cx will be contained in W and so is $cx + y$ if x and y are elements of W. Second let W is a subset satisfying the conditions of this question. Then by picking $a = 1$ or $y = 0$ we get the conditions in Theorem 1.3.

19. It's easy to say that is sufficient since if we have $W_1 \subset W_2$ or $W_2 \subset W_1$ then the union of W_1 and W_2 will be W_1 or W_2, a space of course. To say it's necessary we may assume that neither $W_1 \subset W_2$ nor $W_2 \subset W_1$ holds and then we can find some $x \in W_1 \backslash W_2$ and $y \in W_2 \backslash W_1$. Thus by the condition of subspace we have $x + y$ is a vector in W_1 or in W_2, say W_1. But this will make $y = (x + y) - x$ should be in W_1. It will be contradictory to the original hypothesis that $y \in W_2 \backslash W_1$.

20. We have that $a_i w_i \in W$ for all i. And we can get the conclusion that $a_1 w_1$, $a_1 w_1 + a_2 w_2$, $a_1 w_1 + a_2 w_2 + a_3 w_3$ are in W inductively.

21. In calculus course it will be proven that $\{a_n + b_n\}$ and $\{ca_n\}$ will converge. And zero sequence, that is sequence with all entris zero, will be the zero vector.

22. The fact that it's closed has been proved in the previous exercise. And a zero function is either a even function or odd function.

23. (a) We have $(x_1 + x_2) + (y_1 + y_2) = (x_1 + y_1) + (x_2 + y_2) \in W_1 + W_2$ and $c(x_1 + x_2) = cx_1 + cx_2 \in W_1 + W_2$ if $x_1, y_1 \in W_1$ and $x_2, y_2 \in W_2$. And we have $0 = 0 + 0 \in W_1 + W_2$. Finally $W_1 = \{x + 0 : x \in W_1, 0 \in W_2\} \subset W_1 + W_2$ and it's similar for the case of W_2.

    (b) If U is a subspace contains both W_1 and W_2 then $x + y$ should be a vector in U for all $x \in W_1$ and $y \in W_2$.

24. It's natural that $W_1 \cap W_2 = \{0\}$. And we have $\mathbb{F}^n = \{(a_1, a_2, \ldots, a_n) : a_i \in \mathbb{F}\} = \{(a_1, a_2, \ldots, a_{n-1}, 0) + (0, 0, \ldots, a_n) : a_i \in \mathbb{F}\} = W_1 \oplus W_2$.

25. This is similar to the exercise 1.3.24.

26. This is similar to the exercise 1.3.24.

27. This is similar to the exercise 1.3.24.

28. By the previous exercise we have $(M_1 + M_2)^t = M_1^t + M_2^t = -(M_1 + M_2)$ and $(cM)^t = cM^t = -cM$. With addition that zero matrix is skew-symmetric we have the set of all skew-symmetric matrices is a space. We have $M_{n \times n}(\mathbb{F}) = \{A : A \in M_{n \times n}(\mathbb{F})\} = \{(A + A^t) + (A - A^t) : A \in M_{n \times n}(\mathbb{F})\} = W_1 + W_2$ and $W_1 \cap W_2 = \{0\}$. The final equality is because $A + A^t$ is symmetric and $A - A^t$ is skew-symmetric. If $\mathbb{F}$ is of characteristic 2, we have $W_1 = W_2$.

29. It's easy that $W_1 \cap W_2 = \{0\}$. And we have $M_{n \times n}(\mathbb{F}) = \{A : A \in M_{n \times n}(\mathbb{F})\} = \{(A - B(A)) + B(A) : A \in M_{n \times n}(\mathbb{F})\} = W_1 + W_2$, where $B(A)$ is the matrix with $B_{ij} = B_{ji} = A_{ij}$ if $i \leq j$.

30. If $V = W_1 \oplus W_2$ and some vector $y \in V$ can be represented as $y = x_1 + x_2 = x_1' + x_2'$, where $x_1, x_1' \in W_1$ and $x_2, x_2' \in W_2$, then we have $x_1 - x_1' \in W_1$ and $x_1 - x_1' = x_2 + x_2' \in W_2$. But since $W_1 \cap W_2 = \{0\}$, we have $x_1 = x_1'$ and $x_2 = x_2'$. Conversely, if each vector in V can be uniquely written as $x_1 + x_2$, then $V = W_1 + W_2$. Now if $x \in W_1 \cap W_2$ and $x \neq 0$, then we have that $x = x + 0$ with $x \in W_1$ and $0 \in W_2$ or $x = 0 + x$ with $0 \in W_1$ and $x \in W_2$, a contradiction.

31. (a) If $v + W$ is a space, we have $0 = v + (-v) \in v + W$ and thus $-v \in W$ and $v \in W$. Conversely, if $v \in W$ we have actually $v + W = W$, a space.

    (b) We can proof that $v_1 + W = v_2 + W$ if and only if $(v_1 - v_2) + W = W$. This is because $(-v_1) + (v_1 + W) = \{-v + v + w : w \in W\} = W$ and $(-v_1) + (v_2 + W) = \{-v_1 + v_2 + w : w \in W\} = (-v_1 + v_2) + W$. So if $(v_1 - v_2) + W = W$, a space, then we have $v_1 - v_2 \in W$ by the previous exercise. And if $v_1 - v_2 \in W$ we can conclude that $(v_1 - v_2) + W = W$.

    (c) We have $(v_1 + W) + (v_2 + W) = (v_1 + v_2) + W = (v_1' + v_2') + W = (v_1' + W) + (v_2' + W)$ since by the previous exercise we have $v_1 - v_1' \in W$ and $v_2 - v_2' \in W$ and thus $(v_1 + v_2) - (v_1' + v_2') \in W$. On the other hand, since $v_1 - v_1' \in W$ implies $av_1 - av_1' = a(v_1 - v_1') \in W$, we have $a(v_1 + W) = a(v_1' + W)$.

    (d) It closed because V is closed. The commutativeity and associativity of addition is also because V is commutative and associative. For the zero element we have $(x + W) + W = x + W$. For the inverse element we have $(x + W) + (-x + W) = W$. For the identity element of multiplication we have $1(x + W) = x + W$. The distribution law and combination law are also followed by the original propositions in V.

But there are one more thing should be checked, that is whether it is well-defined. But this is the exercise 1.3.31.(c).

1.4 Linear Combinations and Systems of Linear Equations

1. (a) Yes. Just pick any coeficient to be zero.

    (b) No. By definition it should be $\{0\}$.

    (c) Yes. Every subspaces of which S is a subset contains $\operatorname{span}(S)$ and $\operatorname{span}(S)$ is a subspace.

    (d) No. This action will change the solution of one system of linear equations.

    (e) Yes.

    (f) No. For example, $0x = 3$ has no solution.

2. (a) Original system $\Leftrightarrow \begin{cases} x_1 & - & x_2 & - & 2x_3 & - & x_4 & = & -3 \\ & & & & x_3 & + & 2x_4 & = & 4 \\ & & & & 4x_3 & + & 8x_4 & = & 16 \end{cases}$. So we have solution is $\{(5 + s - 3t, s, 4 - 2t, t) : s, t \in \mathbb{F}\}$.

    (b) $\{(-2, -4, -3)\}$.

    (c) No solution.

    (d) $\{(-16 - 8s, 9 + 3s, s, 2) : s \in \mathbb{F}\}$.

    (e) $\{(-4 + 10s - 3t, 3 - 3s + 2t, r, s, 5) : s, t \in \mathbb{F}\}$.

    (f) $\{(3, 4, -2)\}$.

3. (a) Yes. Solve the equation $x_1(1, 3, 0) + x_2(2, 4, -1) = (-2, 0, 3)$ and we have the solution $(x_1, x_2) = (4, -3)$.

    (b) Yes.

    (c) No.

    (d) No.

    (e) No.

    (f) Yes.

4. (a) Yes.

    (b) No.

    (c) Yes.

    (d) Yes.

    (e) No.

    (f) No.

5. (a) Yes.

    (b) No.

    (c) No.

    (d) Yes.

    (e) Yes.

    (f) No.

    (g) Yes.

    (h) No.

6. For every $(x_1, x_2, x_3) \in \mathbb{F}^3$ we may assume

$$y_1(1,1,0) + y_2(1,0,1) + y_3(0,1,1) = (x_1, x_2, x_3)$$

and solve the system of linear equation. We got $(x_1, x_2, x_3) = \frac{1}{2}(x_1 - x_2 + x_3)(1,1,0) + \frac{1}{2}(x_1 + x_2 - x_3)(1,0,1) + \frac{1}{2}(-x_1 + x_2 + x_3)(0,1,1)$.

7. For every $(x_1, x_2, \ldots x_n) \in \mathbb{F}^n$ we can write $(x_1, x_2, \ldots x_n) = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$.

8. It's similar to exercise 1.4.7.

9. It's similar to exercise 1.4.7.

10. For $x \neq 0$ the statement is the definition of linear combination and the set is a line. For $x = 0$ the both side of the equation is the set of zero vector and the set is the origin.

11. To prove it's sufficient we can use Theorem 1.5 and then we know $W = \text{span}(W)$ is a subspace. To prove it's necessary we can also use Theorem 1.5. Since W is a subspace contains W, we have $\text{span}(W) \subset W$. On the other hand, it's natural that $\text{span}(W) \supset W$.

12. To prove $\text{span}(S_1) \subset \text{span}(S_2)$ we may let $v \in S_1$. Then we can write $v = a_1 x_1 + a_2 x_2 + \cdots + a_3 x_3$ where x_i is an element of S_1 and so is S_2 for all $n = 1, 2, \ldots, n$. But this means v is a linear combination of S_2 and we complete the proof. If $\text{span}(S_1) = V$, we know $\text{span}(S_2)$ is a subspace containing $\text{span}(S_1)$. So it must be V.

13. We prove $\text{span}(S_1 \cup S_2) \subset \text{span}(S_1) + \text{span}(S_2)$ first. For $v \in \text{span}(S_1 \cup S_2)$ we have $v = \sum_{i=1}^{n} a_i x_i + \sum_{j=1}^{m} b_j y_j$ with $x_i \in S_1$ and $y_j \in S_2$. Since the first summation is in $\text{span}(S_1)$ and the second summation is in $\text{span}(S_2)$, we have $v \in \text{span}(S_1) + \text{span}(S_2)$. For the converse, let $u + v \in \text{span}(S_1) + \text{span}(S_2)$ with $u \in \text{span}(S_1)$ and $v \in \text{span}(S_2)$. We can right $u + v = \sum_{i=1}^{n} a_i x_i + \sum_{j=1}^{m} b_j y_j$ with $x_i \in S_1$ and $y_j \in S_2$ and this means $u + v \in \text{span}(S_1 \cup S_2)$.

14. For $v \in \text{span}(S_1 \cap S_2)$ we may write $v = \sum_{i=1}^{n} a_i x_i$ with $x_i \in S_1$ and $x_i \in S_2$. So v is an element of both $\text{span}(S_1)$ and $\text{span}(S_2)$ and hence an element of $\text{span}(S_1) \cap \text{span}(S_2)$. For example we have if $S_1 = S_2 = (1,0)$ then they are the same and if $S_1 = (1,0)$ and $S_2 = (0,1)$ then we have the left hand side is the set of zero vector and the right hand side is the the plane $\mathbb{R}^2$.

15. If we have both $a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = b_1 v_1 + b_2 v_2 + \cdots b_n v_n$ then we have $(a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \cdots + (a_n - b_n)v_n = 0$. By the property we can deduce that $a_i = b_i$ for all i.

16. When W has finite element the statement holds. Otherwise $W - \{v\}$, where $v \in W$ will be a generating set of W. But there are infinitely many $v \in W$.

1.5   Linear Dependence and Linear Independence

1. (a) No. For example, take $S = \{(1,0), (2,0), (0,1)\}$ and then $(0,1)$ is not a linear combination of the other two.

   (b) Yes. It's because $1\vec{0} = \vec{0}$.

   (c) No. It's independent by the remark after Definition of linearly independent.

   (d) No. For example, we have $S = \{(1,0), (2,0), (0,1)\}$ but $\{(1,0), (0,1)\}$ is linearly independent.

   (e) Yes. This is the contrapositive statement of Theorem 1.6.

   (f) Yes. This is the definition.

2. (a) Linearly dependent. We have $-2 \begin{pmatrix} 1 & -3 \\ -2 & 4 \end{pmatrix} = \begin{pmatrix} -2 & 6 \\ 4 & -8 \end{pmatrix}$. So to check the linearly dependency is to find the nontrivial solution of equation $a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0$. And x_1 and x_2 are the two matrices here.

   (b) Linearly independent.

   (c) Linearly independent.

   (d) Linearly dependent.

   (e) Linearly dependent.

   (f) Linearly independent.

   (g) Linearly dependent.

   (h) Linearly independent.

   (i) Linearly independent.

   (j) Linearly dependent.

3. Let $M_1, M_2, \ldots, M_5$ be those matrices. We have $M_1 + M_2 + M_3 - M_4 - M_5 = 0$.

15

4. If $a_1e_1 + a_2e_2 + \cdots + a_ne_n = (a_1, a_2, \ldots, a_n) = 0$, then by comparing the i-th entry of the vector of both side we have $a_1 = a_2 = \cdots = a_n = 0$.

5. It's similar to exercise 1.5.4.

6. it's similar to exercise 1.5.4.

7. Let E_{ij} be the matrix with the only nonzero ij-entry= 1. Then $\{E_{11}, E_{22}\}$ is the generating set.

8. (a) The equation $x_1(1,1,0) + x_2(1,0,1) + x_3(0,1,1) = 0$ has only nontrivial solution when $\mathbb{F} = \mathbb{R}$.

   (b) When $\mathbb{F}$ has characteristic 2, we have $1 + 1 = 0$ and so $(1,1,0) + (1,0,1) + (0,1,1) = (0,0,0)$.

9. It's sufficient since if $u = tv$ for some $t \in \mathbb{F}$ then we have $u - tv = 0$. While it's also necessary since if $au + bv = 0$ for some $a, b \in \mathbb{F}$ with at least one of the two coefficients not zero then we may assume $a \neq 0$ and $u = -\frac{b}{a}v$.

10. Pick $v_1 = (1,1,0)$, $v_2 = (1,0,0)$, $v_3 = (0,1,0)$. And we have that none of the three is a multiple of another and they are dependent since $v_1 - v_2 - v_3 = 0$.

11. Vector in $\text{span}(S)$ are linear combinations of S and they all have different representation by the remark after Definition of linear independent. So there are 2^n representations and so 2^n vectors.

12. Since S_1 is linearly dependent we have finite vectors $x_1, x_2, \ldots, x_n$ in S_1 and so in S_2 such that $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$ is a nontrivial representation. But the nontrivial representation is also a nontrivial representation of S_2. And the Corollary is just the contrapositive statement of the Theorem 1.6.

13. (a) Sufficiency: If $\{u + v, u - v\}$ is linearly independent we have $a(u + v) + b(u - v) = 0$ implies $a = b = 0$. Assuming that $cu + dv = 0$, we can deduce that $\frac{c+d}{2}(u+v) + \frac{c-d}{2}(u-v) = 0$ and hence $\frac{c+d}{2} = \frac{c-d}{2} = 0$. This means $c = d = 0$ if the characteristc is not two. Necessity: If $\{u, v\}$ is linearly independent we have $au + bv = 0$ implies $a = b = 0$. Assuming that $c(u + v) + d(u - v) = 0$, we can deduce that $(c + d)u + (c - d)v = 0$ and hence $c + d = c - d = 0$ and $2c = 2d = 0$. This means $c = d = 0$ if the characteristc is not two.

   (b) Sufficiency: If $au + bv + cw = 0$ we have $\frac{a+b-c}{2}(u + v) + \frac{a-b+c}{2}(u + w) + \frac{-a+b+c}{2}(v + w) = 0$ and hence $a = b = c = 0$. Necessity: If $a(u + v) + b(u + w) + c(v + w) = 0$ we have $(a + b)u + (a + c)v + (b + c)w = 0$ and hence $a = b = c = 0$.

14. Sufficiency: It's natural that 0 is linearly dependent. If v is a linear combination of $u_1, u_2, \ldots, u_n$, say $v = a_1u_1 + a_2u_2 + \cdots a_nu_n$, then $v - a_1u_1 - a_2u_2 - \cdots - a_nu_n = 0$ implies S is linearly dependent. Necessity: If S is

linearly dependent and $S \neq \{0\}$ we have some nontrivial representation $a_0 u_0 + a_1 u_1 + \cdots + a_n u_n = 0$ with at least one of the coefficients is zero, say $a_0 = 0$ without loss the generality. Then we can let $v = u_0 = -\frac{1}{a_0}(a_1 u_1 + a_2 u_2 + \cdots + a_n u_n)$.

15. Sufficiency: If $u_1 = 0$ then S is linearly independent. If

$$u_{k+1} \in \text{span}(\{u_1, u_2, \ldots, u_k\})$$

for some k, say $u_{k+1} = a_1 u_1 + a_2 u_2 + \cdots + a_k u_k$, then we have $a_1 u_1 + a_2 u_2 + \cdots + a_k u_k - u_{k+1} = 0$ is a nontrivial representation. Necessary: If S is linearly dependent, there are some integer k such that there is some nontrivial representation $a_1 u_1 + a_2 u_2 + \cdots + a_k u_k + a_{k+1} u_{k+1} = 0$. Furthermore we may assume that $a_{k+1} \neq 0$ otherwise we may choose less k until that $a_{k+1} \neq 0$. Hence we have $a_{k+1} = -\frac{1}{a_{k+1}}(a_1 u_1 + a_2 u_2 + \cdots + a_k u_k)$ and so $a_{k+1} \in \text{span}(\{u_1, u_2, \ldots, u_k\})$.

16. Sufficiency: We can prove it by contrapositive statement. If S is linearly dependent we can find $a_1 u_1 + a_2 u_2 + \cdots + a_n u_n = 0$. But thus the finite set $\{u_1, u_2, \ldots, u_n\}$ would be a finite subset of S and it's linearly dependent. Necessary: This is the Threorem 1.6.

17. Let $C_1, C_2, \ldots, C_n$ be the columns of M. Let $a_1 C_1 + a_2 C_2 + \cdots + a_n C_n = 0$ then we have $a_n = 0$ by comparing the n-th entry. And inductively we have $a_{n-1} = 0, a_{n-2} = 0, \ldots, a_1 = 0$.

18. It's similar to exercise 1.5.17.

19. We have $a_1 A_1^t + a_2 A_2^t + \cdots + a_k A_k^t = 0$ implies $a_1 A_1 + a_2 A_2 + \cdots + a_k A_k = 0$. Then we have $a_1 = a_2 = \cdots = a_n = 0$.

20. If $\{f, g\}$ is linearly dependent, then we have $f = kg$. But this means $1 = f(0) = kg(0) = k \times 1$ and hence $k = 1$. And $e^r = f(1) = kg(1) = e^s$ means $r = s$.

1.6   Bases and Dimension

1. (a) No. The empty set is its basis.

   (b) Yes. This is the result of Replacement Theorem.

   (c) No. For example, the set of all polynomials has no finite basis.

   (d) No. $\mathbb{R}^2$ has $\{(1, 0), (1, 1)\}$ and $\{(1, 0), (0, 1)\}$ as bases.

   (e) Yes. This is the Corollary after Replacement Theorem.

   (f) No. It's $n + 1$.

   (g) No. It's $m \times n$.

   (h) Yes. This is the Replaceent Theorem.

(i) No. For $S = 1, 2$, a subset of $\mathbb{R}$, then $5 = 1 \times 1 + 2 \times 2 = 3 \times 1 + 1 \times 2$.

(j) Yes. This is Theorem 1.11.

(k) Yes. It's $\{0\}$ and V respectly.

(l) Yes. This is the Corollary 2 after Replacement Theorem.

2. It's enough to check there are 3 vectors and the set is linear independent.

    (a) Yes.

    (b) No.

    (c) Yes.

    (d) Yes.

    (e) No.

3. (a) No.

    (b) Yes.

    (c) Yes.

    (d) Yes.

    (e) No.

4. It's impossible since the dimension of $\mathbf{P}_3(\mathbb{R})$ is four.

5. It's also impossible since the dimension of $\mathbb{R}^3$ is three.

6. Let E_{ij} be the matrix with the only nonzero ij-entry= 1. Then the sets $\{E_{11}, E_{12}, E_{21}, E_{22}\}$, $\{E_{11}+E_{12}, E_{12}, E_{21}, E_{22}\}$, and $\{E_{11}+E_{21}, E_{12}, E_{21}, E_{22}\}$ are bases of the space.

7. We have first $\{u_1, u_2\}$ is linearly independent. And since $u_3 = -4u_1$ and $u_4 = -3u_1 + 7u_2$, we can check that $\{u_1, u_2, u_5\}$ is linearly independent and hence it's a basis.

8. To solve this kind of questions, we can write the vectors into a matrix as

below and do the Gaussian elimintaion.

$$
M = \begin{pmatrix}
2 & -3 & 4 & -5 & 2 \\
-6 & 9 & -12 & 15 & -6 \\
3 & -2 & 7 & -9 & 1 \\
2 & -8 & 2 & -2 & 6 \\
-1 & 1 & 2 & 1 & -3 \\
0 & -3 & -18 & 9 & 12 \\
1 & 0 & -2 & 3 & -2 \\
2 & -1 & 1 & -9 & 7
\end{pmatrix}
$$

$$
\rightsquigarrow \begin{pmatrix}
0 & -3 & 8 & -11 & 6 \\
0 & 0 & 0 & 0 & 0 \\
0 & -2 & 13 & -18 & 7 \\
0 & -8 & 6 & -8 & 10 \\
0 & 1 & 0 & 4 & -5 \\
0 & 1 & 6 & -3 & -4 \\
1 & 0 & -2 & 3 & -2 \\
0 & -1 & 5 & -15 & 11
\end{pmatrix}
$$

$$
\rightsquigarrow \begin{pmatrix}
0 & 0 & 26 & -20 & -6 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 4 & -5 \\
0 & 0 & 54 & -32 & -22 \\
0 & 0 & -6 & -7 & -1 \\
0 & 1 & 6 & -3 & 4 \\
1 & 0 & -2 & 3 & -2 \\
0 & 0 & 11 & -18 & 7
\end{pmatrix}
$$

$$
\rightsquigarrow \begin{pmatrix}
0 & 0 & 0 & -124 & 124 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 4 & -5 \\
0 & 0 & 0 & -248 & 248 \\
0 & 0 & 0 & 31 & -31 \\
0 & 1 & 6 & -3 & 4 \\
1 & 0 & -2 & 3 & -2 \\
0 & 0 & 0 & -62 & 62
\end{pmatrix}
$$

$$
\rightsquigarrow \begin{pmatrix}
0 & 0 & 0 & -124 & 124 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 4 & -5 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 1 & 6 & -3 & 4 \\
1 & 0 & -2 & 3 & -2 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}
$$

And the row with all entries 0 can be omitted[1]. So $\{u_1, u_3, u_6, u_7\}$ would

_____

[1]Which row with all entries is important here. So actually the operation here is not the

be the basis for W (the answer here will not be unique).

9. If $a_1u_1 + a_2u_2 + a_3u_3 + a_4u_4 = (a_1, a_1 + a_2, a_1 + a_2 + a_3, a_1 + a_2 + a_3 + a_4) = 0$ we have $a_1 = 0$ by comparing the first entry and then $a_2 = a_3 = a_4 = 0$. For the second question we can solve $(a_1, a_2, a_3, a_4) = a_1u_1 + (a_2 - a_1)u_2 + (a_3 - a_2)u_3 + (a_4 - a_3)u_4$.

10. The polynomials found by Lagrange interpolation formula would be the answer. It would have the smallest degree since the set of those polynomials of Lagrange interpolation formula is a basis.

   (a) $-4x^2 - x + 8$.
   (b) $-3x + 12$.
   (c) $-x^3 + 2x^2 + 4x - 5$.
   (d) $2x^3 - x^2 - 6x + 15$.

11. If $\{u, v\}$ is a basis then the dimension of V would be two. So it's enough to check both $\{u + v, au\}$ and $\{au, bv\}$ are linearly independent. Assuming $s(u + v) + tau = (s + ta)u + sv = 0$ we have $s + ta = s = 0$ and hence $s = t = 0$. Assuming $sau + tbv = 0$ we have $sa = tb = 0$ and hence $s = t = 0$.

12. If $\{u, v, w\}$ is a basis then the dimension of V would be three. So it's enough to check $\{u + v + w, v + w, w\}$ is lineaerly independent. Assuming $a(u + v + w) + b(v + w) + cw = au + (a + b)v + (a + b + c)w = 0$ we have $a = a + b = a + b + c = 0$ and hence $a = b = c = 0$.

13. We can substract the second equation by the two times of the first equation. And then we have

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

Let $x_3 = s$ and hence $x_2 = s$ and $x_1 = s$. We have the solution would be $\{(s, s, s) = s(1, 1, 1) : s \in \mathbb{R}\}$. And the basis would be $\{(1, 1, 1)\}$.

14. For W_1 we can observe that by setting $a_2 = p$, $a_3 = q$, $a_4 = s$, and $a_5 = t$ we can solve $a_1 = q + s$. So $W_1 = \{(q+s, p, q, s, t) = p(0, 1, 0, 0, 0) + q(1, 0, 1, 0, 0) + s(1, 0, 0, 1, 0) + t(0, 0, 0, 0, 1) : p, q, s, t \in \mathbb{F}^5\}$. And

$$\{(0, 1, 0, 0, 0), (1, 0, 1, 0, 0), (1, 0, 0, 1, 0), (0, 0, 0, 0, 1)\}$$

is the basis. The dimension is four. And similarly for W_2 we may set $a_4 = s$, $a_5 = t$. And then we have $a_1 = -t$, $a_2 = a_3 = a_4 = s$ and

$$W_2 = \{(-t, s, s, s, t) = s(0, 1, 1, 1, 0) + t(-1, 0, 0, 0, 1) : s, t \in \mathbb{F}^5\}$$

. And hence

$$\{(0, 1, 1, 1, 0), (-1, 0, 0, 0, 1)\}$$

is the basis of W_2. The dimension is two.

standard Gaussian elimination since we can not change the order of two row here.

15. Just solve $A_{11}+A_{22}+\cdots+A_{nn} = 0$ and hence $\{E_{ij}\}_{i\neq j}\cup\{E_{ii}-E_{nn}\}_{i=1,2,\ldots n-1}$ is the basis, where $\{E_{ij}\}$ is the standard basis. And the dimension would be $n^2 - 1$.

16. We have $A_{ij} = 0$ for all $i > j$. Hence the basis could be $\{E_{ij}\}_{i\leq j}$ and the dimension is $\frac{n(n+1)}{2}$.

17. We have $A_{ii} = 0$ and $A_{ij} = A_{ji}$. Hence the basis could be $\{E_{ij} - E_{ji}\}_{i<j}$ and the dimension is $\frac{n(n-1)}{2}$.

18. Let e_i be the sequence with the only nonzero i-th term= 1. Then we have $\{e_i\}_{i\geq 0}$ is a basis. To prove it, we have that every sequence is linear combination of the basis since we only discuss the sequence with finite nonzero entries. Furthermore we have for every finite subset of $\{e_i\}_{i\geq 0}$ is linearly independent.

19. If every vector has a unique representation as linear combination of the set β, this means every vector is a linear combination of β. Furthermore, if there are nontrivial representation of 0, then we have there are two representations, say the trivial and nontrivial one, of 0. This is a contradiction.

20. (a) If $S = \varnothing$ or $S = \{0\}$, then we have $V = \{0\}$ and the empty set can generate V. Otherwise we can choose a nonzero vector u_1 in S, and continuing pick u_{k+1} such that $u_{k+1} \notin \text{span}(\{u_1, u_2, \ldots, u_k\})$. The process would teminate before $k > n$ otherwise we can find linearly independent set with size more than n. If it terminates at $k = n$, then we knoew the set is the desired basis. If it terminates at $k < n$, then this means we cannot find any vector to be the vector u_{k+1}. So any vectors in S is a linear combination of $\beta = \{u_1, u_2, \ldots, u_k\}$ and hence β can generate V since S can. But by Replacement Theorem we have $n \leq k$. This is impossible.

    (b) If S has less than n vectors, the process must terminate at $k < n$. It's impossible.

21. Sufficiency: If the vector space V is finite-dimensional, say dim= n, and it contains an infinite linearly independent subset β, then we can pick an independent subset β' of β such that the size of β' is $n + 1$. Pick a basis α with size n. Since α is a basis, it can generate V. By Replacement Theorem we have $n \geq n+1$. It's a contradiction. Necessity: To find the infinite linearly independent subset, we can let S be the infinite-dimensional vector space and do the process in exercise 1.6.20(a). It cannot terminate at any k otherwise we find a linearly independent set generating the space and hence we find a finite basis.

22. The condition would be that $W_1 \subset W_2$. Let α and β be the basis of $W_1 \cap W_2$ and W_1. Since W_1 and W_2 finite-dimensional, we have α and β are bases with finite size. First if W_1 is not a subset of W_2, we have some vector $v \in W_1 \backslash W_2$. But this means that $v \notin \text{span}(\beta)$ and hence $\beta \cup \{v\}$ would be a

independent set with size greater than that of β. So we can conclude that $\dim(W_1 \cap W_2) = \dim(W_1)$. For the converse, if we have $W_1 \subset W_2$, then we have $W_1 \cap W_2 = W_1$ and hence they have the same dimension.

23. Let α and β be the basis of W_1 and W_2. By the definition we have both α and β are bases with finite size.

    (a) The condition is that $v \in W_1$. If $v \notin W_1 = \text{span}(\alpha)$, thus $\alpha \cup \{v\}$ would be a independent set with size greater than α. By Replacement Theorem we have $\dim(W_1) < \dim(W_2)$. For the converse, if $v \in W_1 = \text{span}(\{v_1, v_2, \ldots, v_k\})$, we actually have $W_2 = \text{span}(\{v_1, v_2, \ldots, v_k, v\}) = \text{span}(\{v_1, v_2, \ldots, v_k\}) = W_1$ and hence they have the same dimension.

    (b) Since we have $W_1 \subset W_2$, we have in general we have $\dim(W_1) < \dim(W_2)$.

24. By exercise 1.5.18 we have $\beta = \{f^{(i)}\}_{i=0,1,\ldots,n}$ is independent since they all have different degree. And since $\dim(\mathbf{P}_n(\mathbb{R})) = n+1$ we can conclude that β is a basis and hence it generate the space $\mathbf{P}_n(\mathbb{R})$).

25. It would be $m + n$ since $(\alpha, 0) \cup (0, \beta)$ would be a basis of Z if α and β are the basis of V and W respectly, where $(\alpha, 0) = \{(u, 0) \in Z : u \in V\}$ and $(0, \beta) = \{(0, u) \in Z : u \in W\}$.

26. It would be n since $\{x - a, x^2 - a^2, \ldots, x^n - a^n\}$ is a basis.

27. The dimension of $W_1 \cap \mathbf{P}_n(\mathbb{F})$ and $W_2 \cap \mathbf{P}_n(\mathbb{F})$ are $\lfloor \frac{n+1}{2} \rfloor$ and $\lceil \frac{n+1}{2} \rceil$ respectly since $\{x^i\}$ with $0 \le i \le n$ is an odd number and $\{x^j\}$ with $0 \le j \le n$ is a even number are bases of the two spaces respectly.

28. If α is the basis of V over $\mathbb{R}$, then we have $\alpha \cup i\alpha$ is the basis of V over $\mathbb{R}$, where $i\alpha = \{iv \in V : v \in \alpha\}$.

29. (a) Using the notation of the Hint, if we assume

$$\sum_{i=1}^{k} a_i u_i + \sum_{i=1}^{m} b_i v_i + \sum_{i=1}^{n} c_i w_i = 0$$

, then we have

$$v = \sum_{i=1}^{m} b_i v_i = -\sum_{i=1}^{k} a_i u_i - \sum_{i=1}^{n} c_i w_i$$

is contained in both W_1 and W_2 and hence in $W_1 \cap W_2$. But if $v \neq 0$ and can be express as $u = \sum_{i=1}^{k} a_i' u_i$, then we have $\sum_{i=1}^{m} b_i v_i - \sum_{i=1}^{k} a_i' u_i = 0$. This is contradictory to that $\{u_1, \ldots, v_1, \ldots\}$ is a basis of W_1. Hence we have

$$v = \sum_{i=1}^{m} b_i v_i = -\sum_{i=1}^{k} a_i u_i - \sum_{i=1}^{n} c_i w_i = 0$$

22

, this means $a_i = b_j = c_l = 0$ for all index i, j, and k. So the set $\beta = \{u_1, \ldots, v_1, \ldots, w_1, \ldots\}$ is linearly independent. Furthermore, for every $x + y \in W_1 + W_2$ with $x \in W_1$ and $y \in W_2$ we can find the representation $x = \sum_{i=1}^{k} d_i u_i + \sum_{i=1}^{m} b_i v_i$ and $y = \sum_{i=1}^{k} d'_i u_i + \sum_{i=1}^{n} c_i w_i$. Hence we have

$$x + y = \sum_{i=1}^{k} (d_i + d'_i) u_i + \sum_{i=1}^{m} b_i v_i + \sum_{i=1}^{n} c_i w_i = 0$$

is linear combination of β. Finally we have $\dim(W_1 + W_2) = k + m + n = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$ and hence $W_1 + W_2$ is finite-dimensional.

(b) With the formula in the previous exercise we have

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$$
$$= \dim(W_1) + \dim(W_2)$$

if and only if $\dim(W_1 \cap W_2) = 0$. And $\dim(W_1 \cap W_2) = 0$ if and only if $W_1 \cap W_2 = \{0\}$. And this is the sufficient and necessary condition for $V = W_1 \oplus W_2$.

30. It can be check W_1 and W_2 are subspaces with dimension 3 and 2. We also can find out that $W_1 \cap W_2 = \{\begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} \in V : a, b \in \mathbb{F}\}$ and it has dimension 1. By the formula of the previous exercise, we have that dimension of $W_1 + W_2$ is $2 + 3 - 1 = 4$.

31. (a) This is the conclusion of $W_1 \cap W_2 \subset W_2$.

   (b) By the formula in 1.6.29(a) we have the left hand side

$$= m + n - \dim(W_1 \cap W_2) \leq m + n$$

   since $\dim(W_1 \cap W_2) \geq 0$.

32. (a) Let W_1 be the xy-plane with $m = 2$ and W_2 be the x-axis with $n = 1$. Then we have $W_1 \cap W_2 = W_2$ has dimension 1.

   (b) Let W_1 be the xy-plane with $m = 2$ and W_2 be the z-axis with $n = 1$. Then we have $W_1 + W_2 = \mathbb{R}^3$ has dimension $3 = 2 + 1$.

   (c) Let W_1 be the xy-plane with $m = 2$ and W_2 be the xz-axis with $n = 2$. Then we have $W_1 \cap W_2$ is the x-axis with dimension 1 and $W_1 + W_2$ is $\mathbb{R}^3$ with dimension $3 \neq 2 + 2$.

33. (a) Since $V = W_1 \oplus W_2$ means $W_1 \cap W_2 = \{0\}$ and a basis is linearly independent and so contains no 0, we have $\beta_1 \cap \beta_2 = \varnothing$. And it a special case of exercise 1.6.29(a) that $\beta_1 \cup \beta_2$ is a basis.

(b) Let u_i and v_j are vectors in β_1 and β_2 respectly. If there is a nonzero vector $u \in W_1 \cap W_2$, we can write $u = \sum_{i=1}^{n} a_i u_i = \sum_{j=1}^{m} b_j v_j$. But it impossible since it will cause

$$\sum_{i=1}^{n} a_i u_i - \sum_{j=1}^{m} b_j v_j = 0$$

. On the other hand, for any $v \in V$, we can write $v = \sum_{i=1}^{n} c_i u_i + \sum_{j=1}^{m} d_j v_j \in W_1 + W_2$. For any $x + y \in W_1 + W_2$ with $x \in W_1$ and $y \in W_2$, we have $x = \sum_{i=1}^{n} e_i u_i$ and $y = \sum_{j=1}^{m} f_j v_j$. Thus we have $x + y = \sum_{i=1}^{n} e_i u_i + \sum_{j=1}^{m} f_j v_j \in V$.

34. (a) Let β be the basis of V and α be the basis of W_1. By Replacement Theorem we can extend α to be a basis α' of V such that $\alpha \subset \alpha'$. By the previous exercise and let $W_2 = \mathrm{span}(\alpha' \backslash \alpha)$, we have $V = W_1 \oplus W_2$.

(b) We may set W_2 to be the y-axis and W_2' to be $\{t(1,1) : t \in \mathbb{R}\}$.

35. (a) Since $\{u_1, u_2, \dots, u_n\}$ is a basis, the linear combination of

$$\{u_{k+1}, u_{k+2}, \dots, u_n\}$$

can not be in $\mathrm{span}(\{u_1, u_2, \dots, u_k\}) = W$. This can make sure that

$$\{u_{k+1} + W, u_{k+2} + W, \dots, u_n + W\}$$

is linearly independent by 1.3.31(b). For all $u + W \in V/W$ we can write

$$
\begin{aligned}
u + W &= (a_1 u_1 + a_2 u_2 + \dots + a_n u_n) + W \\
&= (a_1 u_1 + a_2 u_2 + \dots + a_k u_k) + (a_{k+1} u_{k+1} + a_{k+2} u_{k+2} + \dots + a_n u_n) + W \\
&= (a_{k+1} u_{k+1} + a_{k+2} u_{k+2} + \dots + a_n u_n) + W \\
&= a_{k+1}(u_{k+1} + W) + a_{k+2}(u_{k+2} + W) + \dots + a_n(u_n + W)
\end{aligned}
$$

and hence it's a basis.

(b) By preious argument we have $\dim(V/W) = n - k = \dim(V) - \dim(W)$.

1.7   Maximal Linearly Independent Subsets

1. (a) No. For example, the family $\{(0, n)\}_{n \geq 1}$ of open intervals has no maximal element.

(b) No. For example, the family $\{(0, n)\}_{n \geq 1}$ of open intervals in the set real numbers has no maximal element.

(c) No. For example, the two set in this family $\{1, 2, 2, 3\}$ are both maximal element.

(d) Yes. If there are two maximal elements A and B, we have $A \subset B$ or $B \subset A$ since they are in a chain. But no matter $A \subset B$ or $B \subset A$ implies $A = B$ since they are both maximal elements.

(e) Yes. If there are some independent set containing a basis, then the vector in that independent set but not in the basis cannot be a linear combination of the basis.

(f) Yes. It's naturally independent. And if there are some vector can not be a linear combination of a maximal independent set. We can add it to the maximal independent set and have the new set independent. This is contradictory to the maximality.

2. Basis described in 1.6.18 is a infinite linearly independent subset. So the set of convergent sequences is an infinite-dimensional subspace by 1.6.21.

3. Just as the hint, the set $\{\pi, \pi^2, \ldots\}$ is infinite and independent. So V is indinite-dimensional.

4. By Theorem 1.13 we can extend the basis of W to be a basis of V.

5. This is almost the same as the proof of Theorem 1.8 since the definition of linear combination of a infinite subset β is the linear combinations of finite subset of β.

6. Let $\mathscr{F}$ be the family of all linearly independent subsets of S_2 that contain S_1. We may check there are some set containing each member of a chain for all chain of $\mathscr{F}$ just as the proof in Theorem 1.13. So by Maximal principle there is a maximal element β in $\mathscr{F}$. By the maximality we know β can generate S_2 and hence can generate V. With addition of it's independence we know β is a basis.

7. Let $\mathscr{F}$ be the family of all linearly independent subset of β such that union of it and S is independent. Then for each chain $\mathscr{C}$ of $\mathscr{F}$ we may choose U as the union of all the members of $\mathscr{C}$. We should check U is a member of $\mathscr{F}$. So we should check wether $S \cup U$ is independent. But this is easy since if $\sum_{i=1}^{n} a_i v_i + \sum_{j=1}^{m} b_j u_j = 0$ with $v_i \in S$ and $u_j \in U$, say $u_j \in U_j$ where $\{U_j\}$ are a members of $\mathscr{C}$, then we can pick the maximal element, say U_1, of $\{U_j\}$. Thus we have $u_j \in U_1$ for all j. So $S \cup U$ is independent and hence $a_i = b_j = 0$ for all i and j.

Next, by Maximal principle we can find a maximal element S_1 of $\mathscr{C}$. So $S \cup S_1$ is independent. Furthermore by the maximality of S_1 we know that $S \cup S_1$ can generate β and hence can generate V. This means $S \cup S_1$ is a basis for V.

Chapter 2

Linear Transformations and Matrices

2.1 Linear Transformations, Null Spaces, and Ranges

1. (a) Yes. That's the definition.

   (b) No. Consider a map f from $\mathbb{C}$ over $\mathbb{C}$ to $\mathbb{C}$ over $\mathbb{C}$ by letting $f(x+iy) = x$. Then we have $f(x_1 + iy_1 + x_2 + iy_2) = x_1 + x_2$ but $f(iy) = 0 \neq if(y) = iy$.

   (c) No. This is right when T is a linear trasformation but not right in general. For example,

   $$T : \mathbb{R} \to \mathbb{R}$$
   $$x \mapsto x + 1$$

   It's one-to-one but that $T(x) = 0$ means $x = -1$. For the counterexample of converse statement, consider $f(x) = |x|$.

   (d) Yes. We have $T(0_V) = T(0x) = 0T(0_V)_W = 0_W$, for arbitrary $x \in V$.

   (e) No. It is $\dim(V)$. For example, the transformation mapping the real line to $\{0\}$ will be.

   (f) No. We can map a vector to zero.

   (g) Yes. This is the Corollory after Theorem 2.6.

   (h) No. If $x_2 = 2x_1$, then $T(x_2)$ must be $2T(x_1) = 2y_1$.

2. It's a linear transformation since we have

   $$T((a_1, a_2, a_3) + (b_1, b_2, b_3)) = T(a_1 + b_1, a_2 + b_2, a_3 + b_3) = (a_1 + b_1 - a_2 - b_2, 2a_3 + 2b_3)$$

   $$= (a_1 - a_2, 2a_3) + (b_1 - b_2, 2b_3) = T(a_1, a_2, a_3) + T(b_1, b_2, b_3)$$

and
$$T(ca_1, ca_2, ca_3) = (c(a_1 - a_2), 2ca_3) = cT(a_1, a_2, a_3).$$

$N(T) = \{(a_1, a_1, 0)\}$ with basis $\{(1,1,0)\}$; $R(T) = \mathbb{R}^2$ with basis $\{(1,0),(0,1)\}$. Hence T is not one-to-one but onto.

3. Similarly check this is a linear transformation. $N(T) = \{0\}$ with basis $\varnothing$; $R(T) = \{a_1(1,0,2) + a_2(1,0,-1)\}$ with basis $\{(1,0,2),(1,0,-1)\}$. Hence T is one-to-one but not onto.

4. It's a linear transformation. And $N(T) = \{\begin{pmatrix} a_{11} & 2a_{11} & -4a_{11} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}\}$ with basis

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

; $R(T) = \{\begin{pmatrix} s & t \\ 0 & 0 \end{pmatrix}\}$ with basis $\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\}$.

5. It's a linear transformation. And $N(T) = \{0\}$ with basis $\varnothing$; $R(T) = \{ax^3 + b(x^2 + 1) + cx\}$ with basis $\{x^3, x^2 + 1, x\}$. Hence T is one-to-one but not onto.

6. $N(T)$ is the set of all matrix with trace zero. Hence its basis is $\{E_{ij}\}_{i \neq j} \cup \{E_{ii} - E_{nn}\}_{i=1,2,\dots,n-1}$. $R(T) = \mathbb{F}$ with basis 1.

7. For property 1, we have $T(0) = T(0x) = 0T(x) = 0$, where x is an arbitrary element in V. For property 2, if T is linear, then $T(cx+y) = T(cx)+T(y) = cT(x) + T(y)$; if $T(cx + y) = cT(x) + T(y)$, then we may take $c = 1$ or $y = 0$ and conclude that T is linear. For property 3, just take $c = -1$ in property 3. For property 4, if T is linear, then

$$T(\sum_{i=1}^{n} a_i x_i) = T(a_1 x_1) + T(\sum_{i=1}^{n-1} a_i x_i) = \cdots = \sum_{i=1}^{n} T(a_i x_i) = \sum_{i=1}^{n} a_i T(x_i);$$

if the equation holds, just take $n = 2$ and $a_1 = 1$.

8. Just check the two condition of linear transformation.

9. (a) $T(0,0) \neq (0,0)$.
   (b) $T(2(0,1)) = (0,4) \neq 2T(0,1) = (0,2)$.
   (c) $T(2\frac{\pi}{2},0) = (0,0) \neq 2T(\frac{\pi}{2},0) = (2,0)$.
   (d) $T((1,0) + (-1,0)) = (0,0) \neq T(1,0) + T(0,1) = (2,0)$.
   (e) $T(0,0) \neq (0,0)$.

10. We may take $U(a,b) = a(1,4) + b(1,1)$. By Theorem 2.6, the mapping must be $T = U$. Hence we have $T(2,3) = (5,11)$ and T is one-to-one.

11. This is the result of Theorem 2.6 since $\{(1,1),(2,3)\}$ is a basis. And $T(8,11) = T(2(1,1) + 3(2,3)) = 2T(1,1) + 3T(2,3) = (5,-3,16)$.

12. No. We must have $T(-2,0,-6) = -2T(1,0,3) = (2,2) \neq (2,1)$.

13. Let $\sum_{i=0}^{k} a_i v_i = 0$. Then we have $T(\sum_{i=0}^{k} a_i v_i) = \sum_{i=0}^{k} a_i T(v_i) = 0$ and this implies $a_i = 0$ for all i.

14. (a) The sufficiency is due to that if $T(x) = 0$, $\{x\}$ can not be independent and hence $x = 0$. For the necessity, we may assume $\sum a_i T(v_i) = 0$. Thus we have $T(\sum a_i v_i) = 0$. But since T is one-to-one we have $\sum a_i v_i = 0$ and hence $a_i = 0$ for all proper i.

    (b) The sufficiency has been proven in Exercise 2.1.13. But note that S may be an infinite set. And the necessity has been proven in the previous exercise.

    (c) Since T is one-to-one, we have $T(\beta)$ is linear independent by the previous exercise. And since T is onto, we have $R(T) = W$ and hence $\text{span}(T(\beta)) = R(T) = W$.

15. We actually have $T(\sum_{i=0}^{n} a_i x^i) = \sum_{i=0}^{n} \frac{a_i}{i+1} x^{i+1}$. Hence by detailed check we know it's one-to-one. But it's not onto since no function have integral$= 1$.

16. Similar to the previous exercise we have $T(\sum_{i=0}^{n} a_i x^i) = \sum_{i=0}^{n} i a_i x^{i-1}$. It's onto since $T(\sum_{i=0}^{n} \frac{a_i}{i+1} x^{i+1}) = \sum_{i=0}^{n} a_i x^i$. But it's not one-to-one since $T(1) = T(2) = 0$.

17. (a) Because $\text{rank}(T) \leq \dim(V) < \dim(W)$ by Dimension Theorem, we have $R(T) \subsetneq W$.

    (b) Because $\text{nullity}(T) = \dim(V) - \text{rank}(T) \geq \dim(V) - \dim(W) > 0$ by Dimension Theorem, we have $N(T) \neq \{0\}$.

18. Let $T(x,y) = (y,0)$. Then we have $N(T) = R(T) = \{(x,0) : x \in \mathbb{R}\}$.

19. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ and $T(x,y) = (y,x)$ and U is the identity map from $\mathbb{R}^2 \to \mathbb{R}^2$. Then we have $N(T) = N(U) = \{0\}$ and $R(T) = R(U) = \mathbb{R}^2$.

20. To prove $A = T(V_1)$ is a subspace we can check first $T(0) = 0 \in A$. For $y_1, y_2 \in A$, we have for some $x_1, x_2 \in V_1$ such that $T(x_1) = y_1$ and $T(x_2) = y_2$. Hence we have $T(x_1 + x_2) = y_1 + y_2$ and $T(cx_1) = xy_1$. This means both $y_1 + y_2$ and cy_1 are elements of A.

    To prove that $B = \{x \in V : T(x) \in W_1\}$ is a subspace we can check $T(0) = 0 \in W_1$ and hence $0 \in B$. For $x_1, x_2 \in B$, we have $T(x_1), T(x_2) \in W_1$. Hence we have $T(x_1 + x_2) = T(x_1), T(x_2) \in W_1$ and $T(cx_1) = cT(x_1) \in W_1$. This means both $x_1 + x_2$ and cx_1 are elements of B.

21. (a) To prove T is linear we can check

$$T(\sigma_1 + \sigma_2)(n) = \sigma_1(n+1) + \sigma_2(n+1) = T(\sigma_1)(n) + T(\sigma_2)(n)$$

28

and
$$T(c\sigma)(n) = c\sigma(n+1) = cT(\sigma)(n).$$

And it's similar to prove that U is linear.

(b) It's onto since for any σ in V. We may define another sequence τ such that $\tau(0) = 0$ and $\tau(n+1) = \sigma(n)$ for all $n \geq 1$. Then we have $T(\tau) = \sigma$. And it's not one-to-one since we can define a new σ_0 with $\sigma_0(0) = 1$ and $\sigma_0(n) = 0$ for all $n \geq 2$. Thus we have $\sigma_0 \neq 0$ but $T(\sigma_0) = 0$.

(c) If $T(\sigma)(n) = \sigma(n-1) = 0$ for all $n \geq 2$, we have $\sigma(n) = 0$ for all $n \geq 1$. And let σ_0 be the same sequence in the the previous exercise. We cannot find any sequence who maps to it.

22. Let $T(1,0,0) = a$, $T(0,1,0) = b$, and $T(0,0,1) = c$. Then we have

$$T(x,y,z) = xT(1,0,0) + yT(0,1,0) + zT(0,0,1) = ax + by + cz.$$

On the other hand, we have $T(x_1, x_2, \ldots, x_n) = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$ if T is a mapping from $\mathbb{F}^n$ to $\mathbb{F}$. To prove this, just set $T(e_i) = a_i$, where $\{e_i\}$ is the standard of $\mathbb{F}^n$.

For the case that $T : \mathbb{F}^n \to \mathbb{F}$, actually we have

$$T(x_1, x_2, \ldots, x_n) = \left(\sum_{j=1}^{n} a_{1j} x_j, \sum_{j=2}^{n} a_{2j} x_j, \ldots, \sum_{j=m}^{n} a_{mj} x_j \right)$$

. To prove this, we may set $T(e_j) = (a_{1j}, a_{2j}, \ldots, a_{mj})$.

23. With the help of the previous exercise, we have

$$N(T) = \{(x,y,z) : ax + by + cz = 0\}.$$

Hence it's a plane.

24. (a) It will be $T(a,b) = (0,b)$, since $(a,b) = (0,b) + (a,0)$.

(b) It will be $T(a,b) = (0, b-a)$, since $(a,b) = (0, b-a) + (a,a)$.

25. (a) Let W_1 be the xy-plane and W_2 be the z-axis. And $(a,b,c) = (a,b,0) + (0,0,c)$ would be the unique representation of $W_1 \oplus W_2$.

(b) Since $(a,b,c) = (0,0,c) + (a,b,0)$, we have $T(a,b,c) = (0,0,c)$.

(c) Since $(a,b,c) = (a-c,b,0) + (c,0,c)$, we have $T(a,b,c) = (a-c,b,0)$.

26. (a) Since $V = W_1 \oplus W_2$, every vector x have an unique representation $x = x_1 + x_2$ with $x_1 \in W_1$ and $x_2 \in W_2$. So, now we have

$$T(x + cy) = T(x_1 + x_2 + cy_1 + cy_2)$$

$$= T((x_1 + cy_1) + (x_2 + cy_2)) = x_1 + cy_1 = T(x) + cT(y).$$

And hence it's linear.

On the other hand, we have $x = x + 0$ and hence $T(x) = x$ if $x \in W_1$. And if $x \notin W_1$, this means $x = x_1 + x_2$ with $x_2 \neq 0$ and hence we have $T(x) = x_1 \neq x_1 + x_2$.

(b) If $x_1 \in W_1$ then we have $T(x_1 + 0) = x_1 \in R(T)$; and we also have $R(T) \subset W_1$. If $x_2 \in W_2$ then we have $T(x_2) = T(0 + x_2) = 0$ and hence $x_2 \in N(T)$; and if $x \in N(T)$, we have $x = T(x) + x = 0 + x$ and hence $x \in W_2$.

(c) It would be $T(x) = x$ by (a).

(d) It would be $T(x) = 0$.

27. (a) Let $\{v_1, v_2, \ldots, v_k\}$ be a basis for W and we can extend it to a basis $\beta = \{v_1, v_2, \ldots, v_n\}$ of V. Then we may set $W' =$span$(\{v_{k+1}, v_{k+2}, \ldots, v_n\})$. Thus we have $V = W \oplus W'$ and we can define T be the projection on W along W'.

(b) The two projection in Exercise 2.1.24 would be the example.

28. We have $T(0) = 0 \in \{0\}$, $T(x) \in R(T)$, $T(x) = 0 \in N(T)$ if $x \in N(T)$ and hence they are T-invariant.

29. For $x, y \in W$, we have $x + cy \in W$ since it's a subspace and $T(x), T(y) \in W$ since it's T-invariant and finally $T(x + cy) = T(x) + cT(y)$.

30. Since $T(x) \in W$ for all x, we have W is T-invariant. And that $T_W = I_W$ is due to Exercise 2.1.26(a).

31. (a) If $x \in W$, we have $T(x) \in R(T)$ and $T(x) \in W$ since W is T-invariant. But by the definition of direct sum, we have $T(x) \in R(T) \cap W = \{0\}$ and hence $T(x) = 0$.

(b) By Dimension Theorem we have $\dim(N(T)) =\dim(V)-\dim(R(T))$. And since $V = R(T) \oplus W$, we have $\dim W =\dim(V)-\dim(R(T))$. In addition with $W \subset N(T)$ we can say that $W = N(T)$.

(c) Take T be the mapping in Exercise 2.1.21 and $W = \{0\}$. Thus $W \neq N(T) = \{(a_1, 0, 0, \ldots)\}$.

32. We have $N(T_W) \subset W$ since T_W is a mapping from W to W. For $x \in W$ and $x \in N(T_W)$, we have $T_W(x) = T(x) = 0$ and hence $x \in N(T)$. For the converse, if $x \in N(T) \cap W$, we have $x \in W$ and hence $T_W(x) = T(x) = 0$. So we've proven the first statement. For the second statement, we have

$$R(T_W) = \{y \in W : T_W(x) = y, x \in W\} = \{T_W(x) : x \in W\} = \{T(x) : x \in W\}.$$

33. It's natural that $R(T) \supset$span$(\{T(v) : v \in \beta\})$ since all $T(v)$ is in $R(T)$. And for any $y \in R(T)$ we have $y = T(x)$ for some x. But every x is linear

combination of finite many vectors in basis. That is, $x = \sum_{i=1}^{k} a_i v_i$ for some $v_i \in \beta$. So we have

$$y = T(\sum_{i=1}^{k} a_i v_i) = \sum_{i=1}^{k} a_i T(v_i)$$

is an element in $\mathrm{span}(\{T(v) : v \in \beta\})$.

34. Let T be one linear transformation satisfying all conditions. We have that for any $x \in V$ we can write $x = \sum_{i=1}^{k} a_i v_i$ for some $v_i \in \beta$. So by the definition of linear transformation we have $T(x)$ must be

$$T(\sum_{i=1}^{k} a_i v_i) = \sum_{i=1}^{k} a_i T(v_i) = \sum_{i=1}^{k} a_i f(v_i)$$

,a fixed vector in W. So it would be the unique linear transformation.

35. (a) With the hypothesis $V = R(T) + N(T)$, it's sufficient to say that $R(T) \cap N(T) = \{0\}$. But this is easy since

$$\dim(R(T) \cap N(T)) = \dim(R(T)) + \dim(N(T)) - \dim(R(T) + N(T))$$

$$= \dim(R(T)) + \dim(N(T)) - \dim(V) = 0$$

by Dimension Theorem.

(b) Similarly we have

$$\dim(R(T) + N(T)) = \dim(R(T)) + \dim(N(T)) - \dim(R(T) \cap N(T))$$

$$= \dim(R(T)) + \dim(N(T)) - \dim(\{0\}) = \dim(V)$$

by Dimension Theorem. So we have $V = R(T) + N(T)$.

36. (a) In this case we have $R(T) = V$ and $N(T) = \{(a_1, 0, 0, \ldots)\}$. So naturally we have $V = R(T) + N(T)$. But V is a direct sum of them since $R(T) \cap N(T) = N(T) \neq \{0\}$.

(b) Take $T_1 = U$ in the Exercise 2.1.21. Thus we have $R(T_1) = \{(0, a_1, a_2, \ldots)\}$ and $N(T_1) = \{(0, 0, \ldots)\}$. So we have $R(T_1) \cap N(T_1) = \{0\}$ but $R(T_1) + N(T_1) = R(T_1) \neq V$.

37. Let $c = \frac{a}{b} \in \mathbb{Q}$. We have that

$$T(x) = T(\underbrace{\frac{1}{b}x + \frac{1}{b}x + \cdots + \frac{1}{b}x}_{b \text{ times}}) = bT(\frac{1}{b}x)$$

and hence $T(\frac{1}{b}x) = \frac{1}{b}T(x)$. So finally we have

$$T(cx) = T(\frac{a}{b}x) = T(\underbrace{\frac{1}{b}x + \frac{1}{b}x + \cdots + \frac{1}{b}x}_{a \text{ times}})$$

$$= aT(\frac{1}{b}x) = \frac{a}{b}T(x) = cT(x).$$

38. It's additive since

$$T((x_1 + iy_1) + (x_2 + iy_2)) = (x_1 + x_2) - i(y_1 + y_2) = T(x_1 + iy_1) + T(x_2 + iy_2).$$

But it's not linear since $T(i) = -i \neq iT(1) = 0$.

39. It has been proven in the Hint.

40. (a) It's linear since

$$\eta(u + v) = (u + v) + W = (u + W) + (v + W) = \eta(u) + \eta(v)$$

and

$$\eta(cv) = cv + W = c(v + W) = c\eta(v)$$

by the definition in Exercise 1.3.31. And for all element $v + W$ in V/W we have $\eta(v) = v + W$ and hence it's onto. Finally if $\eta(v) = v + W = 0 + W$ we have $v - 0 = v \in W$. Hence $N(\eta) = W$.

(b) Since it's onto we have $R(T) = V/W$. And we also have $N(\eta) = W$. So by Dimension Theorem we have $\dim(V) = \dim(V/W) + \dim(W)$.

(c) They are almost the same but the proof in Exercise 1.6.35 is a special case of proof in Dimension Theorem.

2.2 The Matrix Representation of a Linear Transformation

1. (a) Yes. This is result of Theorem 2.7.

(b) Yes. This is result of Theorem 2.6.

(c) No. It's a $n \times m$ matrix.

(d) Yes. This is Theorem 2.8.

(e) Yes. This is Theorem 2.7.

(f) No. A transformaion of $\mathcal{L}(V, W)$ can not map element in W in general.

2. (a) We have $T(1, 0) = (2, 3, 1) = 2(1, 0, 0) + 3(0, 1, 0) + 1(0, 0, 1)$ and $T(0, 1) = (-1, 4, 0) = -1(1, 0, 0) + 4(0, 1, 0) + 0(0, 0, 1)$. Hence we get

$$[T]_\beta^\gamma = \begin{pmatrix} 2 & -1 \\ 3 & 4 \\ 1 & 0 \end{pmatrix}.$$

(b)

$$[T]_\beta^\gamma = \begin{pmatrix} 2 & 3 & -1 \\ 1 & 0 & 1 \end{pmatrix}.$$

(c)
$$[T]_\beta^\gamma = \begin{pmatrix} 2 & 1 & -3 \end{pmatrix}.$$

(d)
$$[T]_\beta^\gamma = \begin{pmatrix} 0 & 2 & 1 \\ -1 & 4 & 5 \\ 1 & 0 & 1 \end{pmatrix}.$$

(e)
$$[T]_\beta^\gamma = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix}.$$

(f)
$$[T]_\beta^\gamma = \begin{pmatrix} 0 & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & 0 \end{pmatrix}.$$

(g)
$$[T]_\beta^\gamma = \begin{pmatrix} 1 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

3. Since

$$T(1,0) = (1,1,2) = -\frac{1}{3}(1,1,0) + 0(0,1,1) + \frac{2}{3}(2,2,3)$$
$$T(0,1) = (-1,0,1) = -1(1,1,0) + 1(0,1,1) + 0(2,2,3)$$
$$T(1,2) = (-1,1,4) = -\frac{7}{3}(1,1,0) + 2(0,1,1) + \frac{2}{3}(2,2,3)$$
$$T(2,3) = (-1,2,7) = -\frac{11}{3}(1,1,0) + 3(0,1,1) + \frac{4}{3}(2,2,3)$$

we have

$$[T]_\beta^\gamma = \begin{pmatrix} -\frac{1}{3} & -1 \\ 0 & 1 \\ \frac{2}{3} & 0 \end{pmatrix}$$

and

$$[T]_\alpha^\gamma = \begin{pmatrix} -\frac{7}{3} & -\frac{11}{3} \\ 2 & 3 \\ \frac{2}{3} & \frac{4}{3} \end{pmatrix}.$$

33

4. Since

$$T\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 1 + 0x + 0x^2$$

$$T\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 1 + 0x + 1x^2$$

$$T\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 0 + 0x + 0x^2$$

$$T\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 0 + 2x + 0x^2$$

we have

$$[T]_\beta^\gamma = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

5. (a)

$$[T]_\alpha = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

(b)

$$[T]_\beta^\alpha = \begin{pmatrix} 0 & 1 & 0 \\ 2 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

(c)

$$[T]_\beta^\gamma = \begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix}.$$

(d)

$$[A]_\alpha = \begin{pmatrix} 1 \\ -2 \\ 0 \\ 4 \end{pmatrix}.$$

(e)

$$[f(x)]_\beta = \begin{pmatrix} 3 \\ -6 \\ 1 \end{pmatrix}.$$

(f)

$$[a]_\gamma = \begin{pmatrix} a \end{pmatrix}.$$

6. It would be a vector space since all the condition would be true since they are true in V and W. So just check it.

7. If we have $([T]_\beta^\gamma)_{ij} = A_{ij}$, this means

$$T(v_j) = \sum_{i=1}^m A_{ij} w_i.$$

And hence we have

$$aT(v_j) = \sum_{i=1}^m a A_{ij} w_i.$$

and thus $(a[T]_\beta^\gamma)_{ij} = a A_{ij}$.

8. If $\beta = \{v_1, v_2, \ldots, v_n\}$ and $x = \sum_{i=1}^n a_i v_i$ and $y = \sum_{i=1}^n b_i v_i$, then

$$T(x+cy) = T(\sum_{i=1}^n a_i v_i + c \sum_{i=1}^n b_i v_i) = (a_1+cb_1, a_2+cb_2, \ldots, a_n+cb_n) = T(x)+cT(y).$$

9. It would be linear since for $c \in \mathbb{R}$ we have

$$T((x_1+iy_1)+c(x_2+iy_2)) = (x_1+cx_2)+i(x_2+cy_2) = T(x_1+iy_1)+cT(x_2+iy_2).$$

And the matrix would be

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

10. It would be

$$\begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 1 & & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

11. Take $\{v_1, v_2, \ldots, v_k\}$ be a basis of W and extend it to be $\beta = \{v_1, v_2, \ldots, v_n\}$, the basis of V. Since W is T-invariant, we have $T(v_j) = \sum_{i=1}^k a_{ij} v_i$ if $j = 1, 2, \ldots, k$. This means $([T]_\beta)_{ij} = 0$ if $j = 1, 2, \ldots, k$ and $i = k+1, k+2, \ldots, n$.

12. Let $\{v_1, v_2, \ldots, v_k\}$ and $\{v_{k+1}, v_{k+2}, \ldots, v_n\}$ be the basis of W and W' respectly. By Exercise 1.6.33(b), we have $\beta = \{v_1, v_2, \ldots, v_n\}$ is a basis of V. And thus we have

$$[T]_\beta = \begin{pmatrix} I_k & O \\ O & O \end{pmatrix}$$

is a diagonal matrix.

13. Suppose, by contradiction, that $cT = U$ for some c. Since T is not zero mapping, there is some $x \in V$ and some nonzero vector $y \in W$ such that $T(x) = y \neq 0$. But thus we have $y = \frac{1}{c}cy = \frac{1}{c}U(x) = U(\frac{1}{c}x) \in R(U)$. This means $y \in R(T) \cap R(U)$, a contradiction.

14. It can be checked that differentiation is a linear operator. That is, T_i is an element of $\mathcal{L}(V)$ for all i. Now fix some n, and assume $\sum_{i=1}^{n} a_i T_i = 0$. We have $T_i(x^n) = \frac{n!}{(n-i)!} x^{n-i}$ and thus $\{T_i(x^n)\}_{i=1,2,\ldots,n}$ would be an independent set. Since $\sum_{i=1}^{n} a_i T_i(x^n) = 0$, we have $a_i = 0$ for all i.

15. (a) We have zero map is an element in S^0. And for $T, U \in S^0$, we have $(T + cU)(x) = T(x) + cU(x) = 0$ if $x \in S$.

    (b) Let T be an element of S_2^0. We have $T(x) = 0$ if $x \in S_1 \subset S_2$ and hence T is an element of S_1^0.

    (c) Since $V_1 + V_2$ contains both V_1 and V_2, we have $(V_1 + V_2)^0 \subset V_1^0 \cap V_2^0$ by the previous exercise. To prove the converse direction, we may assume that $T \in V_1^0 \cap V_2^0$. Thus we have $T(x) = 0$ if $x \in V_1$ or $x \in V_2$. For $z = u + v \in V_1 + V_2$ with $u \in V_1$ and $v \in V_2$, we have $T(z) = T(u) + T(v) = 0 + 0 = 0$. So T is an element of $(V_1 + V_2)^0$ and hence we have $(V_1 + V_2)^0 \supset V_1^0 \cap V_2^0$.

16. As the process in the proof of Dimension Theorem, we may pick the same basis $\beta = \{v_1, v_2, \ldots v_n\}$ for V. Write $u_{k+1} = T(v_{k+1}$. It has been proven that $\{u_{k+1}, u_{k+2}, \ldots, u_n\}$ is a basis for $R(T)$. Since $\dim(V) = \dim(W)$, we can extend to a basis $\gamma = \{u_1, u_2, \ldots, u_n\}$ for W. Thus we have

$$[T]_\beta^\gamma = \begin{pmatrix} O & O \\ O & I_{n-k} \end{pmatrix}$$

is a diagonal matrix.

2.3 Composition of Linear Transformations and Matrix Multiplication

1. (a) It should be $[UT]_\alpha^\gamma = [T]_\alpha^\beta [U]_\beta^\gamma$.

   (b) Yes. That's Theorem 2.14.

   (c) No. In general β is not a basis for V.

   (d) Yes. That's Theorem 2.12.

   (e) No. It will be true when $\beta = \alpha$.

   (f) No. We have $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^2 = I$.

   (g) No. T is a transformation from V to W but L_A can only be a transformation from $\mathbb{F}^m$ to $\mathbb{F}^n$.

   (h) No. We have $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2 = I$.

   (i) Yes. That's Theorem 2.15.

   (j) Yes. Since $\delta_{ij} = 1$ only when $i = j$, we have $A_{ij} = \delta_i j$.

2. (a)
$$A(2B + 3C) = \begin{pmatrix} 20 & -9 & 18 \\ 5 & 10 & 8 \end{pmatrix}.$$

$$(AB)D = A(BD) = \begin{pmatrix} 29 \\ -26 \end{pmatrix}.$$

(b)
$$A^t = \begin{pmatrix} 2 & -3 & 4 \\ 5 & 1 & 2 \end{pmatrix}.$$

$$A^t B = \begin{pmatrix} 23 & 19 & 0 \\ 26 & -1 & 10 \end{pmatrix}.$$

$$BC^t = \begin{pmatrix} 12 \\ 16 \\ 29 \end{pmatrix}.$$

$$CB = \begin{pmatrix} 27 & 7 & 9 \end{pmatrix}.$$

$$CA = \begin{pmatrix} 20 & 26 \end{pmatrix}.$$

3. (a) We can calculate that $[U]_\beta^\gamma = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$ and $[T]_\beta = \begin{pmatrix} 2 & 3 & 0 \\ 0 & 3 & 6 \\ 0 & 0 & 4 \end{pmatrix}$

and finally
$$[UT]_\beta^\gamma = \begin{pmatrix} 2 & 6 & 6 \\ 0 & 0 & 4 \\ 2 & 0 & -6 \end{pmatrix}.$$

(b) We can calculate $[h(x)]_\beta = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$ and

$$[U(h(x))]_\beta = [U]_\beta^\gamma [h(x)]_\beta \begin{pmatrix} 1 \\ 1 \\ 5 \end{pmatrix}.$$

4. (a) $[T(A)]_\alpha = \begin{pmatrix} 1 \\ -1 \\ 4 \\ 6 \end{pmatrix}.$

(b) $[T(f(x))]_\alpha = \begin{pmatrix} -6 \\ 2 \\ 0 \\ 6 \end{pmatrix}.$

(c) $[T(A)]_\gamma = \begin{pmatrix} 5 \end{pmatrix}.$

(d) $[T(f(x))]_\gamma = \begin{pmatrix} 12 \end{pmatrix}$.

5. **(b)** We have

$$(a(AB))_{ij} = a\sum_{k=1}^{n} A_{ik}B_{kj}$$

$$\sum_{k=1}^{n} aA_{ik}B_{kj} = ((aA)B)_{ij}$$

$$\sum_{k=1}^{n} A_{ik}aB_{kj} = (A(aB))_{ij}$$

(d) We have $[I(v_i)]_\alpha = e_i$ where v_i is the i-th vector of β.

Corollary. We have by Theorem 2.12

$$A(\sum_{i=1}^{k} a_i B_i) = \sum_{i=1}^{k} A(a_i B_i) = \sum_{i=1}^{k} a_i AB_i$$

and

$$(\sum_{i=1}^{k} a_i C_i)A = \sum_{i=1}^{k} (a_i C_i)A = \sum_{i=1}^{k} a_i C_i A.$$

6. We have $(Be_j)_i = \sum_{k=1}^{n} B_{ik}(e_j)_i = B_{ij}$, since $(e_j)_i = 1$ only when $i = j$ and it's 0 otherwise.

7. **(c)** Just check that for all vector $v \in \mathbb{F}^n$ we have $L_{A+B}(v) = (A + B)v = Av + Bv = L_A(v) + L_B(v)$ and $L_{aA}(v) = aA(v) = aL_A(v)$.

   (f) For all vector $v \in \mathbb{F}^n$ we have $L_{I_n}(v) = I_n(v) = v$.

8. In general we may set $T_1, T_2 \in \mathcal{L}(X,Y)$ and $U_1, U_2 \in \mathcal{L}(W,X)$, and $S \in (V,W)$, and thus we have the following statements.

   (a) $T_1(U_1 + U_2) = TU_1 + TU_2$ and $(U_1 + U_2)T = U_1T + U_2T$.
   (b) $T_1(U_1 S) = (T_1 U_1)S$.
   (c) $TI_X = I_Y T = T$.
   (d) $a(T_1 U_1) = (aT_1)U_1 = T_1(aU_1)$ for all scalars a.

   To prove this, just map arbitrary vector in domain by linear transformations and check whether the vectors producted by different transformations meet.

9. Take $A = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ and $U = L_A$ and $T = L_B$.

10. If A is a diagonal matrix then $A_{ij} \neq 0$ only when $i = j$. Hence we have $A_{ij} = \delta_{ij}A_{ij}$. Conversely, if A is not diagonal, we can find $A_{ij} \neq 0$ for some i, j and $i \neq j$. Thus we have $\delta_{ij}A_{ij} = 0 \neq A_{ij}$.

11. If $T^2 = T$ we may pick $y \in R(T)$ and thus we have $y = T(x)$ for some x and $T(y) = T(T(x)) = T^2(x) = 0$. Hence we conclude that $y \in N(T)$. Conversely if we have $R(T) \subset N(T)$, we have $T^2(x) = T(T(x)) = 0$ since $T(x)$ is an element in $R(T)$ and hence in $N(T)$.

12. (a) If UT is injective, we have that $UT(x) = 0$ implies $x = 0$. Thus we have that if $T(x) = 0$ we also have $UT(x) = 0$ and hence $x = 0$. So T is injective. But U may not be injective. For example, pick $U(x, y, z) = (x, y)$, a mapping from $\mathbb{R}^3$ to $\mathbb{R}^2$, and $T(x, y) = (x, y, 0)$, a mapping from $\mathbb{R}^2$ to $\mathbb{R}^3$.

    (b) If UT is surjective, we have that for all $z \in Z$ there is a vector $x \in V$ such that $UT(x) = z$. Thus we have that if for all $z \in Z$ we have $z = U(T(x))$ and hence U is surjective. But T may not surjective. The example in the previous question could also be the example here.

    (c) For all $z \in Z$, we can find $z = U(y)$ for some $y \in W$ since U is surjective and then find $y = T(x)$ for some $x \in V$ since T is surjective. Thus we have $z = UT(x)$ for some x and hence UT is surjective. On the other hand, if $UT(x) = 0$, this means $T(x) = 0$ since U is injective and $x = 0$ since T is injective.

13. It's natural that we have $\text{tr}(A) = \text{tr}(A^t)$ since $A_{ii} = A_{ii}^t$ for all i. On the other hand we have

$$\text{tr}(AB) = \sum_{i=1}^{n} (AB)_{ii} = \sum_{i=1}^{n} \sum_{k=1}^{n} A_{ik} B_{ki}$$
$$= \sum_{k=1}^{n} \sum_{ik=1}^{n} B_{ki} A_{ik} = \sum_{k=1}^{n} (BA)_{kk}$$
$$= \text{tr}(BA).$$

14. (a) We can write

$$Bz = \begin{pmatrix} \sum_{j=1}^{p} a_j B_{1j} \\ \sum_{j=2}^{p} a_j B_{2j} \\ \vdots \\ \sum_{j=1}^{p} a_j B_{nj} \end{pmatrix}$$
$$= \sum_{j=1}^{p} a_j \begin{pmatrix} B_{1j} \\ B_{2j} \\ \vdots \\ B_{nj} \end{pmatrix} = \sum_{j=1}^{p} a_j v_j.$$

    (b) This is the result of Theorem 2.13(b) and the previous exercise.

    (c) This is instant result of the fact that $wA = A^t w^t$.

    (d) This is also because $AB = B^t A^t$.

15. Let v_t be the t-th column vector of A and we have $v_j = \sum_{t \neq j} a_t v_t$. Thus we have $Mv_j = \sum_{t \neq j} a_t Mv_t$. And hence we get the desired result since Mv_t is the column vector of MA.

16. (a) Since we know $R(T)$ is a T-invariant space, we can view T as a mapping from $R(T)$ to $R(T)$ and call this restricted mapping $T|_{R(T)}$. So now we have that

$$\dim R(T) = \operatorname{rank}(T) = \operatorname{rank}(T^2) = \dim(T(T(V)))$$

$$= \dim(T(R(T)) = \operatorname{rank}(T|_{R(T)}).$$

And so the mapping $T|_{R(T)}$ is surjective and hence injective with the help of the fact $R(T)$ is finite dimensional. This also means $N(T|_{R(T)}) = R(T) \cap N(T) = 0$. This complete the proof of the first statement. For the other, it's sufficient to say that $R(T) + N(T) = V$. But this is instant conclusion of the fact that $R(T) + N(T) \subset V$ and that

$$\dim(R(T) + N(T)) = \dim(R(T)) + \dim(N(T)) - \dim(R(T) \cap N(T))$$

$$= \dim(R(T)) + \dim(N(T)) = \dim(V).$$

(b) In general we have $\operatorname{rank}(T^{s+1}) \leq \operatorname{rank}(T^s)$ since the fact $T^{s+1}(V) = T^s(R(T)) \subset T^s(V)$. But the integer $\operatorname{rank}(T^s)$ can only range from 0 to $\dim(V)$. So there must be some integer k such that

$$\operatorname{rank}(T^k) = \operatorname{rank}(T^{k+1}).$$

And this means $T^{k+1}(V) = T^k(V)$ and hence $T^s(V) = T^k(V)$ for all $s \geq k$. Since $2k \geq k$, we can conclude that $\operatorname{rank}(T^k) = \operatorname{rank}(T^{2k})$ and hence we have $V = R(T^k) \oplus N(T^k)$ by the previous exercise.

17. If $T = T^2$, then we have $V = \{y : T(y) = y\} + N(T)$ since $x = T(x) + (x - T(x))$ and we have $T(T(x)) = T(x)$ and $T(x - T(x)) = T(x) - T(x) = 0$. On the other hand, we have if $x \in \{y : T(y) = y\} \cap N(T)$ then we also have $x = T(x) = 0$. So by arguments above we have $V = \{y : T(y) = y\} \oplus N(T)$. Finally we have that T must be the projection on W_1 along W_2 for some W_1 and W_2 such that $W_1 \oplus W_2 = V$.

18. Let A, B, and C be $m \times n$, $n \times p$, and $p \times q$ matrices respectly. Next we want to claim that $(AB)C = A(BC)$ since

$$((AB)C)_{ij} = \sum_{k=1}^{p} (AB)_{ik} C_{kj} = \sum_{k=1}^{p} \left(\sum_{l=1}^{n} A_{il} B_{lk} \right) C_{kj}$$

$$= \sum_{k=1}^{p} \sum_{l=1}^{n} A_{il} B_{lk} C_{kj} = \sum_{l=1}^{n} \sum_{k=1}^{p} A_{il} B_{lk} C_{kj}$$

$$= \sum_{l=1}^{n} A_{il} \left(\sum_{k=1}^{p} B_{lk} C_{kj} \right) = \sum_{l=1}^{n} A_{il} (BC)_{lj} = (A(BC))_{ij}.$$

For the following questions, I would like to prove them in the languague of Graph Theory. So there are some definitions and details in Appendices.

19. Let $G = \mathcal{G}(B)$ be the graph associated to the symmetric matric B. And $(B^3)_{ii}$ is the number of walk of length 3 from i to i. If i is in some clique, then there must be a walk of length 3 from i back to i since a clique must have number of vertex greater than 3. Conversely, if $(B^3)_{ii}$ is greater than zero, this means there is at least one walke of length 3 from i to i, say $i \to j \to k \to i$. Note that i, j, and k should be different vertices since length is 3 and there is no loop. So i, j, and k must be a triangle, this means three vertices adjacent to each others. So i is contained in $\{i, j, k\}$ and so contained in some clique.

20. We can draw the associated digraph and find the cliques as follow:
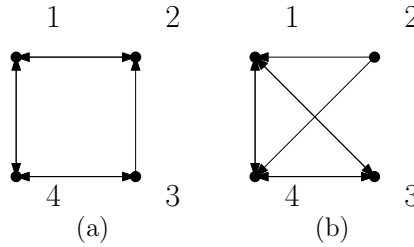
(a) There is no clique.
(b) The only clique would be the set $\{1, 3, 4\}$.

21. A vertex v in a tournament is called a **king** if v can reach all other vertices within two steps. That is, for all vertex u other than v, we have either $v \to u$ or $v \to w \to u$ for some w. So $(A + A^2)_{ij} > 0$ is equivalent to that i can reach j within two steps. And the statement of this question also means that every tournament exists a king.

To prove this statement, we can begin by arbitrary vertex v_1. If v_1 is a king, then we've done. If v_1 is not a king, this means v_1 can not reach some vertex, say v_2, within two steps. Now we have that $d^+(v_2) > d^+(v_1)$ since we have $v_2 \to v_1$ and that if $v_1 \to w$ for some w then we have $v_2 \to w$ otherwise we'll have that $v_1 \to w \to v_2$. Continuing this process and we can find $d^+(v_1) < d^+(v_2) < \cdots$ and terminate at some vertex v_k since there are only finite vertces. And so v_k would be a king.

22. We have $G = \mathcal{G}(A)$ is a tournament drawn below. And every vertex in this tournament could be a king.
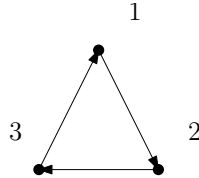
23. The number of nonzero entries would be the number of the edges in a tournament. So it would be $n(n-1)/2$.

2.4   Invertibility and Isomorphisms

1. (a) No. It should be $([T]_\alpha^\beta)^{-1} = [T^{-1}]_\beta^\alpha$.

   (b) Yes. See Appendix B.

   (c) No. L_A can only map $\mathbb{F}^n$ to $\mathbb{F}^m$.

   (d) No. It isomorphic to $\mathbb{F}^5$.

   (e) Yes. This is because $P_n(\mathbb{F}) \cong \mathbb{F}^n$.

   (f) No. We have that $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = I$ but A and B are not invertible since they are not square.

   (g) Yes. Since we have both A and $(A^{-1})^{-1}$ are the inverse of A^{-1}, by the uniqueness of inverse we can conclude that they are the same.

   (h) Yes. We have that $L_{A^{-1}}$ would be the inverse of L_A.

   (i) Yes. This is the definition.

2. (a) No. They have different dimension 2 and 3.

   (b) No. They have different dimension 2 and 3.

   (c) Yes. $T^{-1}(a_1, a_2, a_3) = (-\frac{4}{3}a_2 + \frac{1}{3}a_3, a_2, -\frac{1}{2}a_1 - 2a_2 + \frac{1}{2}a_3)$.

   (d) No. They have different dimension 4 and 3.

   (e) No. They have different dimension 4 and 3.

   (f) Yes. $T^{-1}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} b & a-b \\ c & d-c \end{pmatrix}$.

3. (a) No. They have different dimension 3 and 4.

   (b) Yes. They have the same dimension 4.

   (c) Yes. They have the same dimension 4.

   (d) No. They have different dimension 3 and 4.

4. This is because that $(B^{-1}A^{-1})(AB) = (AB)(B^{-1}A^{-1}) = I$.

5. This is because that $(A^{-1})^t A^t = (AA^{-1})^t = I$ and $A^t(A^{-1})^t = (A^{-1}A)^t = I$.

6. If A is invertible, then A^{-1} exists. So we have $B = A^{-1}AB = A^{-1}O = O/$

7. (a) With the result of the previous exercise, if A is invertible we have that $A = O$. But O is not invertible. So this is a contradiction.

   (b) No. If A is invertible then $B = O$ by the previous exercise.

8. For Corollary 1 we may just pick $W = V$ and $\alpha = \beta$. For Corollary 2 we may just pick $V = \mathbb{F}^n$ and use Corollary 1.

9. If AB is invertible then L_{AB} is invertible. So $L_A L_B = L_{AB}$ is surjective and injective. And thus L_A is injective and L_B surjective by Exercise 2.3.12. But since their domain and codomain has the same dimension, actually they are both invertible, so are A and B.

10. (a) Since $AB = I_n$ is invertible, we have A and B is invertible by the previous exercise.

    (b) We have that $AB = I_n$ and A is invertible. So we can conclude that

    $$A^{-1} = A^{-1} I_n = A^{-1} AB = B.$$

    item Let T is a mapping from V to W and U is a mapping from W to V with dim$W$ =dim$V$. If TU be the identity mapping, then both T and U are invertible. Furthermore $T^{-1} = U$.

    To prove this we may pick bases α of V and β of W and set $A = [T]_\alpha^\beta$ and $B = [U]_\beta^\alpha$. Now apply the above arguments we have that A and B is invertible, so are T and U by Theorem 2.18.

11. If $T(f) = 0$ then we have that $f(1) = f(2) = f(3) = f(4) = 0$, then we have that f is zero function since it has degree at most 3 and it's impossible to have four zeroes if f is nonzero.

12. We can check ϕ_β is linear first. For $x = \sum_{i=1}^n a_i v_i$ and $y = \sum_{i=1}^n b_i v_i$, where

    $$\beta = \{v_1, v_2, \ldots, v_n\}$$

    we have that

    $$\phi_\beta(x + cy) = \begin{pmatrix} a_1 + cb_1 \\ a_2 + cb_2 \\ \vdots \\ a_n + cb_n \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + c \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \phi_\beta(x) + c\phi_\beta(y).$$

    And we can check whether it is injective and surjective. If $\phi_\beta(x) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

    then this means $x = \sum_{i=1}^n 0 v_i = 0$. And for every $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + c$ in $\mathbb{F}^n$, we have

    that $x = \sum_{i=1}^n a_i v_i$ will be associated to it.

13. First we have that V is isomorphic to V by identity mapping. If V is isomorphic to W by mapping T, then T^{-1} exist by the definition of isomorphic and W is isomorphic to V by T^{-1}. If V is isomorphic to W by mapping T and W is isomorphic to X by mapping U, then V is isomorphic to X by mapping UT.

14. Let

$$\beta = \left\{\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$$

be the basis of V. Then we have that ϕ_β in Theorem 2.21 would be the isomorphism.

15. We have that T is isomorphism if and only if that T is injective and surjective. And we also have that the later statement is equivalent to $T(\beta)$ is a baiss for W by Exercise 2.1.14(c).

16. We can check that Φ is linear since

$$\Phi(A + cD) = B^{-1}(A + cD)B = B^{-1}(AB + cDB)$$

$$= B^{-1}AB + cB^{-1}DB = \Phi(A) + c\Phi(D).$$

And it's injective since if $\Phi(A) = B^{-1}AB = O$ then we have $A = BOB^{-1} = O$. It's also be surjective since for each D we have that $\Phi(BDB^{-1}) = D$.

17. (a) If $y_1, y_2 \in T(V_0)$ and $y_1 = T(x_1)$, $y_2 = T(x_2)$, we have that $y_1 + y_2 = T(x_1 + x_2) \in T(V_0)$ and $cy_1 = T(cx_1) = T(V_0)$. Finally since V_0 is a subspace and so $0 = T(0) \in T(V_0)$, $T(V_0)$ is a subspace of W.

(b) We can consider a mapping T' from V_0 to $T(V_0)$ by $T'(x) = T(x)$ for all $x \in V_0$. It's natural that T' is surjective. And it's also injective since T is injective. So by Dimension Theorem we have that

$$\dim(V_0) = \dim(N(T')) + \dim(R(T')) = \dim(T(V_0)).$$

18. With the same notation we have that

$$L_A \phi_\beta(p(x)) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}$$

and

$$\phi_\gamma T(p(x)) = \phi_\gamma(1 + 4x + 3x^2) = \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}.$$

So they are the same.

19. (a) It would be

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

(b) We may check that

$$L_A \phi_\beta \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 2 \\ 4 \end{pmatrix}$$

and

$$\phi_\beta T \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \phi_\beta \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 2 \\ 4 \end{pmatrix}.$$

So they are the same.

20. With the notation in Figure 2.2 we can prove first that $\phi_\gamma(R(T)) = L_A(\mathbb{F}^n)$. Since ϕ_β is surjective we have that

$$L_A(\mathbb{F}^n) = L_A \phi_\beta(V) = \phi_\gamma T(V) = \phi_\gamma(R(T)).$$

Since $R(T)$ is a subspace of W and ϕ_γ is an isomorphism, we have that rank$(T) =$rank(L_A) by Exercise 2.4.17.

On the other hand, we may prove that $\phi_\beta(N(T)) = N(L_A)$. If $y \in \phi_\beta(N(T))$, then we have that $y = \phi_\beta(x)$ for some $x \in N(T)$ and hence

$$L_A(y) = L_A(\phi_\beta(x)) = \phi_\gamma T(x) = \phi_\gamma(0) = 0.$$

Conversely, if $y \in N(L_A)$, then we have that $L_A(y) = 0$. Since ϕ_β is surjective, we have $y = \phi_\beta(x)$ for some $x \in V$. But we also have that

$$\phi_\gamma(T(x)) = L_A(\phi_\beta(x)) = L_A(y) = 0$$

and $T(x) = 0$ since ϕ_γ is injective. So similarly by Exercise 2.4.17 we can conclude that nullity$(T) =$nullity(L_A).

21. First we prove the independence of $\{T_{ij}\}$. Suppose that $\sum_{i,j} a_{ij} T_{ij} = 0$. We have that

$$\left(\sum_{i,j} a_{ij} T_{ij} \right)(v_k) = \sum_i a_{ij} T_{ik}(v_k) = \sum_i a_{ik} w_i = 0.$$

This means $a_{ik} = 0$ for all proper i since $\{w_i\}$ is a basis. And since k is arbitrary we have that $a_{ik} = 0$ for all i and k.

Second we prove that $[T_{ij}]_\beta^\gamma = M^{ij}$. But this is the instant result of

$$T_{ij}(v_j) = w_j$$

and

$$T_{ij}(v_k) = 0$$

for $k \neq j$. Finally we can observe that $\Phi(\beta) = \gamma$ is a basis for $M_{m \times n}(\mathbb{F})$ and so Φ is a isomorphism by Exercise 2.4.15.

45

22. It's linear since

$$T(f + cg) = ((f + cg)(c_0), (f + cg)(c_1), \ldots (f + cg)(c_n))$$

$$= (f(c_0) + cg(c_0), f(c_1) + cg(c_1), \ldots f(c_n) + cg(c_n)) = T(f) + cT(g).$$

Since $T(f) = 0$ means f has $n+1$ zeroes, we know that f must be zero function(This fact can be proven by Lagrange polynomial basis for $P_n(\mathbb{F})$.). So T is injective and it will also be surjective since domain and codomain have same finite dimension.

23. The transformation is linear since

$$T(\sigma + c\tau) = \sum_{i=0}^{m} (\sigma + c\tau)(i)x^i$$

$$= \sum_{i=0}^{m} \sigma(i)x^i + c\tau(i)x^i = T(\sigma) + cT(\tau),$$

where m is a integer large enough such that $\sigma(k) = \tau(k) = 0$ for all $k > m$. It would be injective by following argument. Since $T(\sigma) = \sum_{i=0}^{n} \sigma(i)x^i = 0$ means $\sigma(i) = 0$ for all integer $i \le n$, with the help of the choice of n we can conclude that $\sigma = 0$. On the other hand, it would also be surjective since for all polynomial $\sum_{i=0}^{n} a_i x^i$ we may let $\sigma(i) = a_i$ and thus T will map σ to the polynomial.

24. (a) If $v + N(T) = v' + N(T)$, we have that $v - v' \in N(T)$ and thus $T(v) - T(v') = T(v - v') = 0$.

    (b) We have that

    $$\bar{T}((v + N(T)) + c(u + N(T))) = \bar{T}((v + cu) + N(T))$$

    $$= T(v + cu) = T(v) + cT(u).$$

    (c) Since T is surjective, for all $y \in Z$ we have $y = T(x)$ for some x and hence $y = \bar{T}(x + N(T))$. This means $\bar{T}$ is also surjective. On the other hand, if $\bar{T}(x + N(T)) = T(x) = 0$ then we have that $x \in N(T)$ and hence $x + N(T) = 0 + N(T)$. So $\bar{T}$ is injective. With these argument $\bar{T}$ is an isomorphism.

    (d) For arbitrary $x \in V$, we have

    $$\bar{T}_\eta(x) = \bar{T}(x + N(T)) = T(x).$$

25. The transformation Ψ would be linear since

$$\Psi(f + cg) = \sum_{(f+cg)(s) \ne 0} (f + cg)(s)s = \sum_{(f+cg)(s) \ne 0} f(s)s + cg(s)s$$

$$= \sum_{(f \text{ or } cg)(s) \ne 0} f(s)s + cg(s)s = \sum_{(f \text{ or } cg)(s) \ne 0} f(s) + c \sum_{(f \text{ or } cg)(s) \ne 0} g(s)s$$

$$= \Psi(f) + c\Psi(g).$$

It will be injective by following arguments. If $\Psi(f) = \sum_{f(s)\neq 0} f(s)s = 0$ then we have that $f(s) = 0$ on those s such that $f(s) \neq 0$ since $\{s : f(s) \neq 0\}$ is finite subset of basis. But this can only be possible when $f = 0$. On the other hand, we have for all element $x \in V$ we can write $x = \sum_i a_i s_i$ for some finite subset $\{s_i\}$ of S. Thus we may pick a function f sucht that $f(s_i) = a_i$ for all i and vanish outside. Thus Ψ will map f to x. So Ψ is surjective. And thus it's an isomorphism.

2.5   The Change of Coordinate Matrix

1. (a) No. It should be $[x'_j]_\beta$.

   (b) Yes. This is Theorem 2.22.

   (c) Yes. This is Theorem 2.23.

   (d) No. It should be $B = Q^{-1}AQ$.

   (e) Yes. This is the instant result of the definition of similar and Theorem 2.23.

2. For these problem, just calculate $[I]^{\beta}_{\beta'}$.

   (a) $\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$.

   (b) $\begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$.

   (c) $\begin{pmatrix} 3 & -1 \\ 5 & -2 \end{pmatrix}$.

   (d) $\begin{pmatrix} 2 & -1 \\ 5 & -4 \end{pmatrix}$.

3. (a) $\begin{pmatrix} a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 \\ a_0 & b_0 & c_0 \end{pmatrix}$.

   (b) $\begin{pmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}$.

   (c) $\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ -3 & 2 & 1 \end{pmatrix}$.

   (d) $\begin{pmatrix} 2 & 1 & 1 \\ 3 & -2 & 1 \\ -1 & 3 & 1 \end{pmatrix}$.

(e) $\begin{pmatrix} 5 & -6 & 3 \\ 0 & 4 & -1 \\ 3 & -1 & 2 \end{pmatrix}$.

(f) $\begin{pmatrix} -2 & 1 & 2 \\ 3 & 4 & 1 \\ -1 & 5 & 2 \end{pmatrix}$.

4. We have that
$$[T]_\beta = \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix}$$

and
$$[T]_{\beta'} = [I]_\beta^{\beta'} [[T]_\beta [I]_{\beta'}^\beta$$

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

$$= \begin{pmatrix} 8 & 13 \\ -5 & 9 \end{pmatrix}.$$

5. We have that
$$[T]_\beta = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

and
$$[T]_{\beta'} = [I]_\beta^{\beta'} [[T]_\beta [I]_{\beta'}^\beta$$

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

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

6. Let α be the standard basis (of $\mathbb{F}^2$ or $\mathbb{F}^3$). We have that $A = [L_A]_\alpha$ and hence $[L_A]_\beta = [I]_\alpha^\beta [L_A]_\alpha [I]_\beta^\alpha$. So now we can calculate $[L_A]_\beta$ and $Q = [I]_\beta^\alpha$ and $Q^{-1} = [I]_\alpha^\beta$.

(a) $[L_A]_\beta = \begin{pmatrix} 6 & 11 \\ -2 & -4 \end{pmatrix}$ and $Q = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$.

(b) $[L_A]_\beta = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$ and $Q = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

(c) $[L_A]_\beta = \begin{pmatrix} 2 & 2 & 2 \\ -2 & -3 & -4 \\ 1 & 1 & 2 \end{pmatrix}$ and $Q = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}$.

(d) $[L_A]_\beta = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 18 \end{pmatrix}$ and $Q = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ -2 & 0 & 1 \end{pmatrix}$.

48

7. We may let β be the standard basis and $\alpha = \{(1, m), (-m, 1)\}$ be another basis for $\mathbb{R}^2$.

(a) We have that $[T]_\alpha = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and $Q^{-1} = [I]_\alpha^\beta = \begin{pmatrix} 1 & m \\ -m & 1 \end{pmatrix}$. We also can calculate that $Q = [I]_\beta^\alpha = \begin{pmatrix} \frac{1}{m^2+1} & \frac{m}{m^2+1} \\ -\frac{m}{m^2+1} & \frac{1}{m^2+1} \end{pmatrix}$. So finally we get

$$[T]_\beta = Q^{-1}[T]_\alpha Q = \begin{pmatrix} \frac{1-m^2}{m^2+1} & \frac{2m}{m^2+1} \\ \frac{2m}{m^2+1} & \frac{m^2-1}{m^2+1} \end{pmatrix}.$$

That is, $T(x, y) = \left(\frac{x+2ym-xm^2}{m^2+1}, \frac{-y+2xm+ym^2}{m^2+1} \right)$.

(b) Similarly we have that $[T]_\alpha = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. And with the same Q and Q^{-1} we get

$$[T]_\beta = Q^{-1}[T]_\alpha Q = \begin{pmatrix} \frac{1}{m^2+1} & \frac{m}{m^2+1} \\ \frac{m}{m^2+1} & \frac{m^2}{m^2+1} \end{pmatrix}.$$

That is, $T(x, y) = \left(\frac{x+ym}{m^2+1}, \frac{xm+ym^2}{m^2+1} \right)$.

8. This is similar to the proof of Theorem 2.23 since

$$[T]_{\beta'}^{\gamma'} = [I_W]_\gamma^{\gamma'}[T]_\beta^\gamma[I_V]_{\beta'}^\beta = P^{-1}[T]_\beta^\gamma Q.$$

9. We may denote that A is similar to B by $A \sim B$. First we have $A = I^{-1}AI$ and hence $A \sim A$. Second, if $A \sim B$ we have $A = Q^{-1}BQ$ and $B = (Q^{-1})^{-1}AQ^{-1}$ and hence $B \sim A$. Finally, if $A \sim B$ and $B \sim C$ then we have $A = P^{-1}BP$ and $B = Q^{-1}CQ$. And this means $A = (QP)^{-1}C(QP)$ and hence $A \sim C$. So it's a equivalence relation.

10. If A and B are similar, we have $A = Q^{-1}BQ$ for some invertible matrix Q. So we have
$$\text{tr}(A) = \text{tr}(Q^{-1}BQ) = \text{tr}(QQ^{-1}B) = \text{tr}(B)$$

by Exercise 2.3.13.

11. (a) This is because
$$RQ = [I]_\beta^\gamma[I]_\alpha^\beta = [I]_\alpha^\gamma.$$

(b) This is because
$$Q^{-1} = ([I]_\alpha^\beta)^{-1} = [I^{-1}]_\beta^\alpha = [I]_\beta^\alpha.$$

12. This is the instant result that $A = [L_A]_\beta$ and Q defined in the Corollary is actually $[I]_\gamma^\beta$.

13. Since Q is invertible, we have that L_Q is invertible. We try to check β' is an independent set and hence a basis since V has dimension n. Suppose that $\sum_{j=1}^{n} a_j x'_j = 0$. And it means that

$$\sum_{j=1}^{n} a_j \sum_{i=1}^{n} Q_{ij} x_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_j Q_{ij} \right) x_i = 0.$$

Since β is a basis, we have that $\sum_{j=1}^{n} a_j Q_{ij} = 0$ for all i. Actually this is a system of linear equations and can be written as

$$\begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} \begin{pmatrix} Q_{11} & Q_{12} & \cdots & Q_{1n} \\ Q_{21} & Q_{22} & \cdots & Q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{n1} & Q_{n2} & \cdots & Q_{nn} \end{pmatrix} = vQ = 0,$$

where $v = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}$. But since Q is invertible and so Q^{-1} exist, we can deduce that $v = vQQ^{-1} = 0Q^{-1} = 0$. So we know that β is a basis. And it's easy to see that $Q = [I]_\beta^\beta$ is the change of coordinate matrix changing β'-coordinates into β-coordinates.

14. Let $V = \mathbb{F}^n$, $W = \mathbb{F}^m$, $T = L_A$, β, and γ be the notation defined in the Hint. Let β' and γ' be the set of column vectors of Q and P respectly. By Exercise 2.5.13 we have that β' and γ' are bases and $Q = [I]_{\beta'}^\beta$, $P = [I]_{\gamma'}^\gamma$. Since we have that $[T]_{\beta'}^{\gamma'} = [I]_\gamma^{\gamma'} [T]_\beta^\gamma [I]_{\beta'}^\beta$, we have that $B = P^{-1} A Q$.

2.6   Dual Spaces

1. (a) No. Every linear functional is a linear transformation.

   (b) Yes. It's domain and codomain has dimension 1.

   (c) Yes. They have the same dimension.

   (d) Yes. It's isomorphic to the dual space of its dual space. But if the "is" here in this question means "equal", then it may not be true since dual space must has that its codomain should be $\mathbb{F}$.

   (e) No. For an easy example we may let T be the linear transformation such that $T(x_i) = 2f_i$, where $\beta\{x_1, x_2, \dots, x_n\}$ is the basis for V and $\beta^*\{f_1, f_2, \dots, f_n\}$ is the corresponding dual basis for V^*.

   (f) Yes.

   (g) Yes. They have the same dimension.

   (h) No. Codomain of a linear functional should be the field.

2. In these question we should check whether it's linear and whether its domain and codomain are V and F respectly.

(a) Yes. We may check that

$$f(p(x) + cq(x)) = 2p'(0) + 2cp'(0) + p''(1) + cq''(1)$$

$$2p'(0) + p''(1) + c(2p'(0) + q''(1)) = f(p(x)) + cf(q(x)).$$

(b) No. It's codomain should be the field.

(c) Yes. We may check that

$$\operatorname{tr}(A + cB) = \sum_{i=1}^{n}(A + cB)_{ii}$$

$$= \sum_{i=1}^{n} A_{ii} + cB_{ii} = \operatorname{tr}(A) + c\operatorname{tr}(B).$$

(d) No. It's not linear.

(e) Yes. We may check that

$$f(p(x) + cq(x)) = \int_{0}^{1}(p(t) + cq(t))dt$$

$$= \int_{0}^{1} p(t)dt + c\int_{0}^{1} q(t)dt = f(p(x)) + cf(q(x)).$$

(f) Yes. We may check that

$$f(A + cB) = (A + cB)_{11} = A_{11} + cB_{11} = f(A) + cf(B).$$

3. (a) We may find out that for all vector $(x, y, z) \in \mathbb{R}^3$ we can express it as

$$(x, y, z) = (x - \frac{y}{2})(1, 0, 1) + \frac{y}{2}(1, 2, 1) + (z - x)(0, 0, 1).$$

So we can write
$$\begin{cases} f_1(x, y, z) = x - \frac{y}{2}; \\ f_2(x, y, z) = \frac{y}{2}; \\ f_3(x, y, z) = z - x. \end{cases}$$

(b) This is much easier and we have that
$$\begin{cases} f_1(a_0 + a_1 x + a_2 x^2) = a_0; \\ f_2(a_0 + a_1 x + a_2 x^2) = a_1; \\ f_3(a_0 + a_1 x + a_2 x^2) = a_2. \end{cases}$$

4. We may returen the representation such that

$$(x, y, z) = (x - 2y)(\frac{2}{5}, -\frac{3}{10}, -\frac{1}{10}) + (x + y + z)(\frac{3}{5}, \frac{3}{10}, \frac{1}{10}) + (y - 3z)(\frac{1}{5}, \frac{1}{10}, -\frac{3}{10}).$$

We may check that the set $\{(\frac{2}{5}, -\frac{3}{10}, -\frac{1}{10}), (\frac{3}{5}, \frac{3}{10}, \frac{1}{10}), (\frac{1}{5}, \frac{1}{10}, -\frac{3}{10})\}$ is a basis and hence the desired set. By Theorem 2.24 $\{f_1, f_2, f_3\}$ is a basis for V^*.

5. Assume that $p(t) = a + bx$. We have $\int_0^1 (a + bt)dt = a + \frac{b}{2}$ and $\int_0^2 (a + bt)dt = 2a + 2b$. So we may returen the representation such that

$$a + bx = (a + \frac{b}{2})(2 - 2x) + (2a + 2b)(-\frac{1}{2} + x).$$

We may check that the set $\{2 - 2x, -\frac{1}{2} + x\}$ is a basis and hence the desired set. By Theorem 2.24 $\{f_1, f_2, f_3\}$ is a basis for V^*.

6. (a) Calculate directly that

$$T^t(f)(x, y) = fT(x, y) = f(3x + 2y, x) = 7x + 4y.$$

   (b) Since $\beta = \{(1, 0), (0, 1)\}$ and $(x, y) = x(1, 0) + y(0, 1)$, we have that $f_1(x, y) = x$ and $f_2(x, y) = y$. So we can find out that

$$T^t(f_1)(x, y) = f_1T(x, y) = f_1(3x + 2y, x) = 3x + 2y = 3f_1(x, y) + 2f_2(x, y);$$

$$T^t(f_2)(x, y) = f_2T(x, y) = f_2(3x + 2y, x) = x = 1f_1(x, y) + 0f_2(x, y).$$

And we have the matrix $[T^t]_{\beta^*} = \begin{pmatrix} 3 & 1 \\ 2 & 0 \end{pmatrix}$.

   (c) Since $T(x, y) = (3x + 2y, x)$, we can calculate that

$$[T]_\beta = \begin{pmatrix} 3 & 2 \\ 1 & 0 \end{pmatrix}$$

and

$$([T]_\beta)^t = \begin{pmatrix} 3 & 1 \\ 2 & 0 \end{pmatrix}.$$

So we have that $[T^t]_{\beta^*} = ([T]_\beta)^t$.

7. (a) Calculate directly that

$$T^t(f)(a + bx) = fT(a + bx) = f(-a - 2b, a + b) = -3a - 4b.$$

   (b) Since $\beta = \{1, x\}$ and $a + bx = a \times 1 + b \times x$, we have that $f_1(a + bx) = a$ and $f_2(a + bx) = b$. And since $\gamma = \{(1, 0), (0, 1)\}$ and $(a, b) = a(1, 0) + b(0, 1)$, we have that $g_1(a, b) = a$ and $g_2(a, b) = b$. So we can find out that

$$T^t(g_1)(a + bx) = g_1T(a + bx) = g_1(-a - 2b, a + b) = -a - 2b$$

$$= -1 \times g_1(a, b) + (-2) \times g_2(a, b);$$

$$T^t(g_2)(a + bx) = g_2T(a + bx) = g_2(-a - 2b, a + b) = a + b$$

$$= 1 \times g_1(a, b) + 1 \times g_2(a, b).$$

And we have the matrix $[T^t]_{\gamma^*}^{\beta^*} = \begin{pmatrix} -1 & 1 \\ -2 & 1 \end{pmatrix}$.

(c) Since $T(a + bx) = (-a - 2b, a + b)$, we can calculate that

$$[T]_\beta^\gamma = \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}$$

and

$$([T]_\beta^\gamma)^t = \begin{pmatrix} -1 & 1 \\ -2 & 1 \end{pmatrix}.$$

So we have that $[T^t]_{\gamma^*}^{\beta^*} = ([T]_\beta^\gamma)^t$.

8. Every plane could be written in the form $P = \{(x, y, z) : ax+by+cz = 0\}$ for some scalar a, b and c. Consider a transformation $T(x, y, z) = ax + by + cz$. It can be shown that T is an element in $(\mathbb{R}^3)^*$ and $P = N(T)$. For the case in $\mathbb{R}^2$, actually every line has the form $L = \{(x, y) : ax + by = 0\}$ and hence is the null space of a vector in $(\mathbb{R}^2)^*$.

9. If T is linear, we can set f_i be g_iT as the Hint. Since it's conposition of two linear function, it's linear. So we have

$$T(x) = (g_1(T(x)), g_2(T(x)), \ldots, g_m(T(x)))$$

$$= (f_1(x), f_2(x), \ldots, f_m(x)).$$

For the converse, let $\{e_i\}_{i=1,2,\ldots,m}$ be the standard basis of $\mathbb{F}^m$. So if we have that $T(x) = \sum_{i=1}^m f_i(x)e_i$ with f_i linear, we can define $T_i(x) = f_i(x)e_i$ and it would be a linear transformation in $\mathscr{L}(\mathbb{F}^n, \mathbb{F}^m)$. Thus we know T is linear since T is summation of all T_i.

10. (a) Since we can check that $f_i(p(x) + cq(x)) = p(c_i) + cq(c_i) = f_i(p(x)) + cf_i(q(x))$, f_i is linear and hence in V^*. And we know that $\dim(V^*) = \dim(V) = \dim(P_n(\mathbb{F})) = n + 1$. So now it's enough to show that $\{f_0, f_1, \ldots, f_n\}$ is independent. So assume that $\sum_{i=1}^n a_i f_i = 0$ for some a_i. We may define polynomials $p_i(x) = \prod_{j \ne i} (x - c_j)$ such that we know $p_i(c_i) \ne 0$ but $p_i(c_j) = 0$ for all $j \ne i$. So now we have that

$$\sum_{i=1}^n a_i f_i(p_1) = a_1 f_1(p_1) = 0$$

implies $a_1 = 0$. Similarly we have $a_i = 0$ for all proper i.

(b) By the Corollary after Theorem 2.26 we have an ordered basis

$$\beta = \{p_0, p_1, \ldots, p_n\}$$

for V such that $\{f_1, f_2, \ldots, f_n\}$ defined in the previous exercise is its dual basis. So we know that $p_i(c_j) = \delta_{ij}$. Since β is a basis, every polynomial in V is linear combination of β. If a polynomial q has the

property that $q(c_j) = \delta_{0j}$, we can assume that $q = \sum_{i=0}^{n} a_i p_i$. Then we have

$$1 = q(c_0) = \sum_{i=0}^{n} a_i p_i(c_0) = a_1$$

and

$$0 = q(c_j) = \sum_{i=0}^{n} a_i p_i(c_j) = a_j$$

for all j other than 1. So actually we know $q = p_0$. This means p_0 is unique. And similarly we know all p_i is unique. Since the Lagrange polynomials,say $\{r_i\}_{i=1,2,\ldots n}$, defined in Section 1.6 satisfy the property $r_i(c_j) = \delta_{ij}$, by uniqueness we have $r_i = p_i$ for all i.

(c) Let $\beta = \{p_0, p_1, \ldots, p_n\}$ be those polynomials defined above. We may check that

$$q(x) = \sum_{i=0}^{n} a_i p_i(x)$$

has the property $q(c_i) = a_i$ for all i, since we know that $p_i(c_j) = \delta_{ij}$. Next if $r(x) \in V$ also has the property, we may assume that

$$r(x) = \sum_{i=0}^{n} b_i p_i(x)$$

since β is a basis for V. Similarly we have that

$$a_i = r(c_i) = \sum_{i=0}^{n} b_i p_i(c_i) = b_i.$$

So we know $r = q$ and q is unique.

(d) This is the instant result of 2.6.10(a) and 2.6.10(b) by setting $a_i = p(c_i)$.

(e) Since there are only finite term in that summation, we have that the order of integration and summation can be changed. So we know

$$\int_a^b p(t)dt = \int_a^b \left(\sum_{i=0}^{n} p(c_i)p_i(t)\right)dt$$

$$= \sum_{i=0}^{n} \int_a^b p(c_i)p_i(t)dt.$$

11. It will be more clearer that we confirm that the domain and codomain of both $\psi_2 T$ and $T^{tt}\psi_2$ are V and W^{**} respectively first. So for all $x \in V$ we have

$$\psi_2 T(x) = \psi(T(x)) = \hat{T(x)} \in W^{**}$$

and

$$T^{tt}\psi_1(x) = T^{tt}(\hat{x})$$

$$= (T^t)^t(\hat{x}) = \hat{x}T^t \in W^{**}.$$

But to determine whether two elements f and g in W^{**} are the same is to check whether the value of $f(h)$ and $g(h)$ are the same for all $h \in W^*$. So let h be an element in $W*$. Let's check that

$$\hat{T(x)}(h) = h(T(x))$$

and

$$\hat{x}T^t(h) = \hat{x}(hT) = h(T(x)).$$

So we know they are the same.

12. Let $\beta = \{x_1, x_2, \ldots, x_n\}$ be a basis for V. Then we know the functional $\hat{x}_i \in V^{**}$ means $\hat{x}_i(f) = f(x_i)$ for all funtional f in V^*. On the other hand, we have the dual basis $\beta^* = \{f_1, f_2, \ldots, f_n\}$ is defined by $f_i(x_j) = \delta_{ij}$ for all $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, n$ such that f_i is lineaer. And we can further ferret what elements are in β^{**}. By definition of β^{**} we know $\beta^{**} = \{F_1, F_2, \ldots, F_n\}$ and $F_i(f_j) = \delta_{ij}$ and F_i is linear. So we may check that whether $F_i = \hat{x}_i$ by

$$\hat{x}_i(f_j) = f_j(x_i) = \delta_{ij} = F_i(f_j).$$

Since they are all linear functional and the value of them meets at basis β, they are actually equal by the Corollary after Theorem 2.6.

13. (a) We can check that $f + g$ and cf are elements in S^0 if f and g are elements in S^0 since $(f + g)(x) = f(x) + g(x) = 0$ and $(cf)(x) = cf(x) = 0$. And the zero function is an element in S^0.

(b) Let $\{v_1, v_2, \ldots, v_k\}$ be the basis of W. Since $x \notin W$ we know that $\{v_1, v_2, \ldots, v_{k+1} = x\}$ is an independent set and hence we can extend it to a basis $\{v_1, v_2, \ldots, v_n\}$ for V. So we can define a linear transformation T such that $f(v_i) = \delta_{i(k+1)}$. And thus f is the desired functional.

(c) Let W be the subspace span(S). We first prove that $W^0 = S^0$. Since every function who is zero at W must be a function who is zero at S. we know $W^0 \subset S^0$. On the other hand, if a linear function has the property that $f(x) = 0$ for all $x \in S$, we can deduce that $f(y) = 0$ for all $y \in W =$span(S). Hence we know that $W^0 \supset S^0$ and $W^0 = S^0$. Since $(W^0)^0 = (S^0)^0$ and span$(\psi(S)) = \psi(W)$ by the fact ψ is an isomorphism, we can just prove that $(W^0)^0 = \psi(W)$.

Next, by Theorem 2.26 we may assume every element in $(W^0)^0 \subset V^{**}$ has the form $\hat{x}$ for some x. Let $\hat{x}$ is an element in $(W^0)^0$. We have that $\hat{x}(f) = f(x) = 0$ if $f \in W^0$. Now if x is not an element in W, by the previous exercise there exist some functional $f \in W^0$ such that $f(x) \neq 0$. But this is a contradiction. So we know that $\hat{x}$ is an element in $\psi(W)$ and $(W^0)^0 \subset \psi(W)$.

For the converse, we may assume that $\hat{x}$ is an element in $\psi(W)$. Thus for all $f \in W^0$ we have that $\hat{x}(f) = f(x) = 0$ since x is an element in W. So we know that $(W^0)^0 \supset \psi(W)$ and get the desired conclusion.

(d) It's natural that if $W_1 = W_2$ then we have $W_1^0 = W_2^0$. For the converse, if $W_1^0 = W_2^0$ then we have

$$\psi(W_1) = (W_1^0)^0 = (W_2^0)^0 = \psi(W_2)$$

and hence
$$W_1 = \psi^{-1}\psi(W_1) = \psi^{-1}\psi(W_2) = W_2$$

by the fact that ψ is an isomorphism.

(e) If f is an element in $(W_1 + W_2)^0$, we have that $f(w_1 + w_2) = 0$ for all $w_1 \in W_1$ and $w_2 \in W_2$. So we know that $f(w_1 + 0) = 0$ and $f(0 + w_2) = 0$ for all proper w_1 and w_2. This means f is an element in $W_1^0 \cap W_2^0$. For the converse, if f is an element in $W_1^0 \cap W_2^0$, we have that $f(w_1 + w_2) = f(w_1) + f(w_2) = 0$ for all $w_1 \in W_1$ and $w_2 \in W_2$. Hence we have that f is an element in $(W_1 + W_2)^0$.

14. We use the notation in the Hint. To prove that $\alpha = \{f_{k+1}, f_{k+2}, \ldots, f_n\}$ is a basis for W^0, we should only need to prove that $\operatorname{span}(\alpha) = W^0$ since by $\alpha \subset \beta^*$ we already know that α is an independent set. Since $W^0 \subset V^*$, every element $f \in W^0$ we could write $f = \sum_{i=1}^{n} a_i f_i$. Next since for $1 \le i \le k$ x_i is an element in W, we know that

$$0 = f(x_i) = \sum_{i=1}^{n} a_i f_i(x_i) = a_i.$$

So actually we have $f = \sum_{i=k+1}^{n} a_i f_i$ is an element in $\operatorname{span}(\alpha)$. And finally we get the conclusion by

$$\dim(W) + \dim(W_0) = k + (n - k) = n = \dim(V).$$

15. If $T^t(f) = fT = 0$, this means $f(y) = 0$ for all $y \in R(T)$ and hence $f \in (R(T))^0$. If $f \in (R(T))^0$, this means $f(y) = 0$ for all $y \in R(T)$ and hence $T^t(f)(x) = f(T(x)) = 0$ for all x. This means f is an element in $N(T^t)$.

16. We have that

$$\operatorname{rank}(L_A) = \dim(R(L_A)) = m - \dim(R(L_A)^0) = m - \dim(N((L_A)^t))$$

$$= \dim((\mathbb{F}^m)^*) - \dim(N((L_A)^t)) = \dim(R((L_A)^t)).$$

Next, let α, β be the standard basis for $\mathbb{F}^n$ and $\mathbb{F}^m$. Let α^*, β^* be their dual basis. So we have that $[(L_A)^t]_{\beta^*}^{\alpha^*} = ([L_A]_\alpha^\beta)^t = A^t$ by Theorem 2.25. Let ϕ_{β^*} be the isomorphism defined in Theorem 2.21. We get

$$\dim(R((L_A)^t)) = \dim(\phi_{\beta^*}(R((L_A)^t))) = \dim(R(L_{A^t})) = \operatorname{rank}(L_{A^t}).$$

17. If W is T-invariant, we have that $T(W) \subset W$. Let f be a functional in W^0. We can check $T^t(f) = fT$ is an element in W^0 since $T(w) \in W$ by the fact that T-invariant and thus $f(T(w)) = 0$.

    For the converse, if W^0 is T^t-invariant, we know $T^t(W^0) \subset W^0$. Fix one w in W, if $T(w)$ is not an element in W, by Exercise 2.6.13(b) there exist a functional $f \in W^0$ such that $f(T(w)) \ne 0$. But this means $T^t(f)(w) = fT(w) \ne 0$ and hence $T^t(f) \notin W^0$. This is a contradiction. So we know that $T(w)$ is an element in W for all w in W.

18. First check that Φ is a linear transformation by
    $$\Phi(f + cg)(s) = (f + cg)_S(s) = f_S(s) + cg_S(s) = (\Phi(f) + c\Phi(g))(s).$$

    Second we know Φ is injective and surjective by Exercise 2.1.34.

19. Let S' is a basis for W and we can extend it to be a basis S for V. Since W is a proper subspace of V, we have at least one element $t \in S$ sucht that $t \notin W$. And we can define a function g in $\mathcal{F}(S, \mathbb{F})$ by $g(t) = 1$ and $g(s) = 0$ for all $s \in S$. By the previous exercise we know there is one unique linear functional $f \in V^*$ such that $f_S = g$. Finally since $f(s) = 0$ for all $s \in S'$ we have $f(s) = 0$ for all $s \in W$ but $f(t) = 1$. So f is the desired functional.

20. (a) Assume that T is surjective. We may check whether $N(T^t) = \{0\}$ or not. If $T^t(f) = fT = 0$, we have that $f(y) = f(T(x)) = 0$ for all $y \in W$ since there exist some $x \in V$ such that $T(x) = y$. For the converse, assume that T^t is injective. Suppose, by contradiction, $R(T) \ne W$. By the previous exercise we can construct a nonzero linear functional $f(y) \in W^*$ such that $f(y) = 0$ for all $y \in R(T)$. Let f_0 be the zero functional in W^*. But now we have that $T^t(f)(x) = f(T(x)) = 0 = T^t(g)(x)$, a contradiction. So T must be surjective.

    (b) Assume that T^t is surjective. Suppose, by contradiction, $T(x) = 0$ for some nonzero $x \in V$. We can construct a nonzero linear functional $g \in V^*$ such that $g(x) \ne 0$. Since T^t is surjective, we get some functional $f \in W^*$ such that $T^t(f) = g$. But this means
    $$0 = f(T(x)) = T^t(f)(x) = g(x) \ne 0,$$
    a contradiction.

    For the converse, assume that T is injective and let S is a basis for V. Since T is injective, we have $T(S)$ is an independent set in W. So we can extend it to be a basis S' for W. Thus for every linear functional $g \in V^*$ we can construct a functional $f \in W^*$ such that $T^t(f) = g$ by the argument below. First we can construct a function $h \in \mathcal{F}(S, \mathbb{F})$ by $h(T(s)) = g(s)$ for $s \in S$ and $h(t) = 0$ for all $t \in S' \backslash T(S)$. By Exercise 2.6.18 there is a lineaer functional $f \in W^*$ such that $f_{S'} = h$. So now we have for all $s \in S$
    $$g(s) = h(T(s)) = f(T(s)) = T^t(f)(s).$$

    By Exercise 2.1.34 we have $g = T^t(f)$ and get the desired conclusion.

2.7 Homogeneous Linear Differential Equations with Constant Coeficients

1. (a) Yes. It comes from Theorem 2.32.

   (b) Yes. It comes from Theorem 2.28.

   (c) No. The equation $y = 0$ has the auxiliary polynomial $p(t) = 1$. But $y = 1$ is not a solution.

   (d) No. The function $y = e^t + e^{-t}$ is a solution to the linear differential equation $y'' - y = 0$.

   (e) Yes. The differential operator is linear.

   (f) No. The differential equation $y'' - 2y' + y = 0$ has a solution space of dimension two. So $\{e^t\}$ could not be a basis.

   (g) Yes. Just pick the differential equation $p(D)(y) = 0$.

2. (a) No. Let W be a finite-dimensional subspace generated by the function $y = t$. Thus y is a solution to the trivial equation $0y = 0$. But the solution space is C^∞ but not W. Since $y^{(k)} = 0$ for $k \geq 2$ and it is impossible that

   $$ay' + by = a + bt = 0$$

   for nonzero a, W cannot be the solution space of a homogeneous linear differential equation with constant coefficients.

   (b) No. By the previous argument, the solution subspace containing $y = t$ must be C^∞.

   (c) Yes. If x is a solution to the homogeneous linear differential equation with constant coefficients whose is auxiliary polynomial $p(t)$, then we can compute that $p(D)(x') = D(p(D)(x)) = 0$.

   (d) Yes. Compute that

   $$p(D)q(D)(x + y) = q(D)p(D)(x) + p(D)q(D)(x) = 0.$$

   (e) No. For example, e^t is a solution for $y' - y = 0$ and e^{-t} is a solution for $y' + y = 0$, but $1 = e^t e^{-t}$ is not a solution for $y'' - y = 0$.

3. Use Theorem 2.34.

   (a) The basis is $\{e^{-t}, te^{-t}\}$.

   (b) The basis is $\{1, e^t, e^{-t}\}$.

   (c) The basis is $\{e^t, te^t, e^{-t}, te^{-t}\}$.

   (d) The basis is $\{e^{-t}, te^{-t}\}$.

   (e) The basis is $\{e^{-t}, e^{\alpha t}, e^{\bar{\alpha}t}\}$, where α is the complex value $1 + 2i$.

4. Use Theorem 2.34.

(a) The basis is $\{e^{\alpha t}, e^{\beta t}\}$, where $\alpha = \frac{1+\sqrt{5}}{2}$ and $\beta = \frac{1-\sqrt{5}}{2}$.

(b) The basis is $\{e^t, te^t, t^2 e^t\}$.

(c) The basis is $\{1, e^{-2t}, e^{-4t}\}$.

5. If f and g are elements in C^∞, then we know that the k-th derivative of $f + g$ exists for all integer k since

$$(f + g)^{(k)} = f^{(k)} + g^{(k)}.$$

So $f + g$ is also an element in C^∞. Similarly, for any scalar c, the k-th derivative of cf exists for all integer k since $(cf)^{(k)} = cf^{(k)}$. Finally, the function $f = 0$ is an element in C^∞ naturally.

6. (a) Use the fact
$$D(f + cg) = D(f) + cD(g)$$

for functions $f, g \in C^\infty$ and scalar c. This fact is a easy property given in the Calculus course.

(b) If $p(t)$ is a polynomial, then the differential operator $p(D)$ is linear by Theorem E.3.

7. Let W and V be the two subspaces generated by the two sets $\{x, y\}$ and $\{\frac{1}{2}(x+y), \frac{1}{2i}(x-y)\}$ separately. We know that $W \supset V$ since $\frac{1}{2}(x+y)$ and $\frac{1}{2i}(x-y)$ are elements in W. And it is also true that $W \subset V$ since

$$x = \frac{1}{2}(x + y) + \frac{i}{2i}(x - y)$$

and

$$y = \frac{1}{2}(x + y) - \frac{i}{2i}(x - y)$$

are elements in V.

8. Compute that $e^{(a\pm ib)t} = e^{at} e^{ibt} = (\cos bt + i \sin bt)e^{at}$. By Theorem 2.34 and the previous exercise we get the result.

9. Since those U_i are pairwise commutative, we may just assume that $i = n$. Hence if $U_n(x) = 0$ for some $x \in V$, then

$$U_1 U_2 \cdots U_n(x) = U_1 U_2 \cdots U_{n-1}(0) = 0.$$

10. Use induction on the number n of distinct scalar c_i's. When $n = 1$, the set $\{e^{c_1 t}\}$ is independent since $e^{c_1 t}$ is not identically zero. Suppose now the set $\{e^{c_1 t}, e^{c_2 t}, \ldots, e^{c_n t}\}$ is independent for all $n < k$ and for distinct c_i's. Assume that

$$\sum_{i=1}^{k} b_i e^{c_i t} = 0.$$

Since any differential operator is linear, we have

$$0 = (D - c_k I)(\sum_{i=1}^{k} b_i e^{c_k t}) = \sum_{i=1}^{k-1} (c_i - c_k) b_i e^{c_i t}.$$

This means that $(c_i - c_k) b_i = 0$ and so $b_i = 0$ for all $i < k$ by the fact that c_i's are all distinct. Finally b_k is also zero since

$$b_k e^{c_k t} = 0.$$

11. Denote the given set in Theorem 2.34 to be S. All the element in the set S is a solution by the proof of the Lemma before Theorem 2.34. Next, we prove that S is linearly independent by induction on the number k of distinct zeroes. For the case $k = 1$, it has been proven by the Lemma before Theorem 2.34. Suppose now the set S is linearly independent for the case $k < m$. Assume that

$$\sum_{i=1}^{m} \sum_{j=0}^{n_i-1} b_{i,j} t^j e^{c_i t} = 0$$

for some coefficient $b_{i,j}$. Observe that

$$(D - c_m I)(t^j e^{c_i t}) = j t^{j-1} e^{c_i t} + (c_i - c_m) t^j e^{c_i t}.$$

Since any differential operator is linear, we have

$$(D - c_m I)^{n_m} (\sum_{i=1}^{m} \sum_{j=0}^{n_i-1} b_{i,j} t^j e^{c_i t}) = 0.$$

Since all terms fo $i = m$ are vanished by the differential operator, we may apply the induction hypothesis and know the coefficients for all terms in the left and side is zero. Observer that the coefficient of the term $t^{n_i-1} e^{c_i t}$ is $(c_i - c_m)^{n_m} b_{i,n_i-1}$. This means $(c_i - c_m)^{n_m} b_{i,n_i-1} = 0$ and so $b_{i,n_i-1} = 0$ for all $i < m$. Thus we know that the coefficient of the term $t^{n_i-2} e^{c_i t}$ is $(c_i - c_m)^{n_m} b_{i,n_i-2}$. Hence $b_{i,n_i-2} = 0$ for all $i < m$. Doing this inductively, we get $b_{i,j} = 0$ for all $i < m$. Finally, the equality

$$\sum_{j=0}^{n_m-1} b_{m,j} t^j e^{c_m t} = 0$$

implies $b_{m,j} = 0$ for all j by the Lemma before Theorem 2.34. Thus we complete the proof.

12. The second equality is the definition of range. To prove the first equality, we observe that $R(g(D_V)) \subset N(h(D))$ since

$$h(D)(g(D)(V)) = p(D)(V) = \{0\}.$$

Next observe that
$$N(g(D_V)) = N(g(D))$$
since $N(g(D))$ is a subspace in V. By Theorem 2.32, the dimension of $N(g(D_V)) = N(g(D))$ is the degree of g. So the dimension of $R(g(D_V))$ is the degree of $h(t)$ minus the degree of $g(t)$, that is the degree of $h(t)$. So $N(h(D))$ and $R(g(D_V))$ have the same dimension. Hence they are the same.

13. (a) The equation could be rewriten as $p(D)(y) = x$, where $p(t)$ is the auxiliary polynomial of the equation. Since D is surjective by the Lemma 1 after Theorem 2.32, the differential operator $p(D)$ is also surjective. Hence we may find some solution y_0 such that $p(D)(y_0) = x$.

(b) Use the same notation in the previous question. We already know that $p(D)(z) = x$. If w is also a solution such that $p(D)(w) = x$, then we have

$$p(D)(w - z) = p(D)(w) - p(D)(z) = x - x = 0.$$

So all the solution must be of the form $z + y$ for some y in the solution space V for the homogeneous linear equation.

14. We use induction on the order n of the equation. Let $p(t)$ be the auxiliary polynomial of the equation. If now $p(t) = t - c$ for some coeficient c, then the solution is Ce^{ct} for some constant C by Theorem 2.34. So if $Ce^{ct_0} = 0$ for some $t_0 \in \mathbb{R}$, then we know that $C = 0$ and the solution is the zero function. Suppose the statement is true for $n < k$. Now assume the degree of $p(t)$ is k. Let x be a solution and t_0 is a real number. For an arbitrary scalar c, we factor $p(t) = q(t)(t - c)$ for a polynomial $q(t)$ of degree $k - 1$ and set $z = q(D)(x)$. We have $(D - cI)(z) = 0$ since x is a solution and $z(t_0) = 0$ since $x^{(i)}(t_0) = 0$ for all $0 \le i \le n-1$. Again, z must be of the form Ce^{ct}. And so $Ce^{ct_0} = 0$ implies $C = 0$. Thus z is the zero function. Now we have $q(D)(x) = z = 0$. By induction hypothesis, we get the conclusion that x is identically zero. This complete the proof.

15. (a) The mapping Φ is linear since the differential operator D is linear. If $\Phi(x) = 0$, then x is the zero function by the previouse exercise. Hence Φ is injective. And the solution space is an n-dimensional space by Theorem 2.32. So The mapping is an isomorphism.

(b) This comes from the fact the transformation Φ defined in the previous question is an isomorphism.

16. (a) Use Theorem 2.34. The auxiliary polynomial is $t^2 + \frac{g}{l}$. Hence the basis of the solution space is

$$\{e^{it\sqrt{\frac{g}{l}}}, e^{-it\sqrt{\frac{g}{l}}}\}$$

or
$$\{\cos t\sqrt{\frac{g}{l}}, \sin t\sqrt{\frac{g}{l}}\}$$

by Exercise 2.7.8. So the solution should be of the form

$$\theta(t) = C_1 \cos t\sqrt{\frac{g}{l}} + C_2 \sin t\sqrt{\frac{g}{l}}$$

for some constants C_1 and C_2.

(b) Assume that

$$\theta(t) = C_1 \cos t\sqrt{\frac{g}{l}} + C_2 \sin t\sqrt{\frac{g}{l}}$$

for some constants C_1 and C_2 by the previous argument. Consider the two initial conditions

$$\theta(0) = C_1\sqrt{\frac{g}{l}} = \theta_0$$

and

$$\theta'(0) = C_2\sqrt{\frac{g}{j}} = 0.$$

Thus we get

$$C_1 = \theta_0\sqrt{\frac{l}{g}}$$

and

$$C_2 = 0.$$

So we get the unique solution

$$\theta(t) = \theta_0\sqrt{\frac{l}{g}}\cos t\sqrt{\frac{g}{l}}.$$

(c) The period of $\cos t\sqrt{\frac{g}{l}}$ is $2\pi\sqrt{\frac{l}{g}}$. Since the solution is unique by the previous argument, the pendulum also has the same period.

17. The auxiliary polynomial is $t^2 + \frac{k}{m}$. So the general solution is

$$y(t) = C_1 \cos t\sqrt{\frac{k}{m}} + C_2 \sin t\sqrt{\frac{k}{m}}$$

for some constants C_1 and C_2 by Exercise 2.7.8.

18. (a) The auxiliary polynomial is $mt^2 + rt + k$. The polynomial has two zeroes

$$\alpha = \frac{-r + \sqrt{r^2 - 4mk}}{2m}$$

and
$$\beta = \frac{-r - \sqrt{r^2 - 4mk}}{2m}.$$

So the general solution to the equation is

$$y(t) = C_1 e^{\alpha t} + C_2 e^{\beta t}.$$

(b) By the previous argument assume the solution is

$$y(t) = C_1 e^{\alpha t} + C_2 e^{\beta t}.$$

Consider the two initial conditions

$$y(0) = C_1 + C_2 = 0$$

and

$$y'(0) = \alpha C_1 + \beta C_2 = v_0.$$

Solve that

$$C_1 = (\alpha - \beta)^{-1} v_0$$

and

$$C_2 = (\beta - \alpha)^{-1} v_0.$$

(c) The limit tends to zero since the real parts of α and β is both $-\frac{r}{2m}$, a negative value, by assuming the $r^2 - 4mk \le 0$. Even if $r^2 - 4mk > 0$, we still know that α and β are negative real number.

19. Since $\mathcal{F}(\mathbb{R}, \mathbb{R})$ is a subset of $\mathcal{F}(\mathbb{C}, \mathbb{C})$, so if the solution which is useful in describing physical motion, then it will still be a real-valued function.

20. (a) Assume the differential equation has monic auxiliary polynomial $p(t)$ of degree n. Thus we know that $p(D)(x) = 0$ if x is a solution. This means that $x^{(k)}$ exists for all integer $k \le n$. We may write $p(t)$ as $t^n + q(t)$, where $q(t) = p(t) - t^n$ is a polynomial of degree less than n. Thus we have
$$x^{(n)} = -q(D)(x)$$

is differentiable since $x^{(n)}$ is a linear combination of lower order terms $x^{(k)}$ with $k \le n - 1$. Doing this inductively, we know actualy x is an element in C^∞.

(b) For complex number c and d, we may write $c = c_1 + ic_2$ and $d = d_1 + id_2$ for some real numbers c_1, c_2, d_1, and d_2. Thus we have

$$e^{c+d} = e^{(c_1+d_1)+i(c_2+d_2)} = e^{c_1} e^{d_1} \left(\cos(c_2 + d_2) + i \sin(c_2 + d_2) \right)$$

and
$$e^c e^d = e_1^c e_1^d (\cos c_2 + i \sin c_2)(\cos d_2 + i \sin d_2)$$
$$= e_1^c e_1^d [(\cos c_2 \cos d_2 - \sin c_2 \sin d_2) + i(\sin c_2 \cos d_2 + \cos c_2 \sin d_2)]$$

$$= e^{c_1}e^{d_1}(\cos(c_2 + d_2) + i\sin(c_2 + d_2)).$$

This means $e^{c+d} = e^c e^d$ even if c and d are complex number.[1] For the second equality, we have

$$1 = e^0 = e^{c-c} = e^c e^{-c}.$$

So we get

$$e^{-c} = \frac{1}{e^c}.$$

(c) Let V be the set of all solution to the homogeneous linear differential equation with constant coefficient with auxiliary polynomial $p(t)$. Since each solution is an element in C^∞, we know that $V \supset N(p(D))$, where $N(p(D))$ is the null space of $p(D)$, since $p(D)(x) = 0$ means that x is a solution. Conversely, if x is a solution, then we have $p(D)(x) = 0$ and so $x \in N(p(D))$.

(d) Let $c = c_1 + ic_2$ for some real numbers c_1 and c_2. Directly compute that

$$(e^{ct})' = (e^{c_1 t + ic_2 t})' = (e^{c_1 t}(\cos c_2 t + i\sin c_2 t))'$$

$$c_1 e^{c_1 t}(\cos c_2 t + i\sin c_2 t)) + ic_2 e^{c_1 t}(\cos c_2 t + i\sin c_2 t)$$

$$(c_1 + ic_2)e^{c_1 t}(\cos c_2 t + i\sin c_2 t) = ce^{ct}.$$

(e) Assume that $x = x_1 + ix_2$ and $y = y_1 + iy_2$ for some x_1, x_2, y_1, and y_2 in $\mathcal{F}(\mathbb{R}, \mathbb{R})$. Compute that

$$(xy)' = (x_1 y_1 - x_2 y_2)' + i(x_1 y_2 + x_2 y_1)'$$

$$= (x_1' y_1 + x_1 y_1' - x_2' y_2 - x_2 y_2') + i(x_1' y_2 + x_1 y_2' + x_2' y_1 + x_2 y_1')$$

$$= (x_1' + ix_2')(y_1 + iy_2) + (x_1 + ix_2)(y_1' + iy_2') = x'y + xy'.$$

(f) Assume that $x = x_1 + ix_2$ for some x_1 and x_2 in $\mathcal{F}(\mathbb{R}, \mathbb{R})$. If

$$x' = x_1' + ix_2' = 0,$$

then $x_1' = 0$ and $x_2' = 0$ since x_1' and x_2' are real-valued functions. Hence x_1 and x_2 are constant in $\mathbb{R}$. Hence x is a constant in $\mathbb{C}$.

---

[1] The textbook has a typo that $e^{c+d} = c^c e^d$.

Chapter 3

Elementary Matrix Operations and Systems of Linear Equations

3.1 Elementary Matrix Operations and Elementary Matrices

1. (a) Yes. Since every elementary matrix comes from I_n, a square matrix.

    (b) No. For example, $2I_1$ is an elementary matrix of type 2.

    (c) Yes. It's an elementary matrix of type 2 with scalar 1.

    (d) No. For example, the product of two elementary matrices

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

    is not an elementary matrix.

    (e) Yes. This is Theorem 3.2.

    (f) No. For example, the sum of two elementary matrices

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

    is not an elementary matrix.

    (g) Yes. See Exercise 3.1.5.

    (h) No. For example, let $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$. Then we can obtain B by add one time the first row of A to the second row of B. But all column operation on A can not change the fact that the second row of A is two zeros.

(i) Yes. If $B = EA$, we have $E^{-1}B = A$ and E^{-1} is an elementary matrix of row operation.

2. By adding -2 times the first column of A to the second column, we obtain B. By adding -1 time the first row of B to the second row, we obtain C. Finally let $E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$, $E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix}$,

$E_4 = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $E_5 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$. We have that

$$E_5 E_4 E_3 E_2 E_1 C = I_3.$$

The following is the process.

$$C = \begin{pmatrix} 1 & 0 & 3 \\ 0 & -2 & -2 \\ 1 & -3 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 1 \\ 1 & -3 & 1 \end{pmatrix}$$

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

$$\rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

3. (a) This matrix interchanges the first and the third row. So the inverse matrix do the inverse step. So the inverse matrix do the same thing and it is

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

(b) This matrix multiplies the second row by 3. To the inverse matrix multiplies the second row by $\frac{1}{3}$ and it is

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

(c) This matrix adds -2 times the first row to the third row. So the invers matrix adds 2 times the first row to the third row and it is

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

4. A matrix who interchanges the i-th and the j-th rows is also a matrix who interchanges the i-th and the j-th columns. A matrix who multiplies the i-th row by scalar c is also a matrix who multiplies the i-th column by scalar c. A matrix who adds c times the i-th row to the j-th row is also a matrix who adds c times the j-th column to the i-th column.

5. We can check that matrices of type 1 or type 2 are symmetric. And the transpose of a matrix of type 3, who adds c times the i-th row(column) to the j-th row(column), is a matrix of type 3, who adds c times the j-th row(column) to the i-th row(column).

6. If B can be obtained from A by an elementary row operation, we could write $B = EA$. So we have $B^t = A^t E^t$ and this means B can be obtained by A by elementary column operation with corresponding elementary matrix E^t. If B can be obtained from A by an elementary column operation, we could write $B = AE$. So we have $B^t = E^t A^t$ and this means B can be obtained by A by elementary row operation with corresponding elementary matrix E^t.

7. It's enough to check the following matrix multiplication is right. Let $\{u_1, u_2, \ldots, u_n\}$ and $\{v_1, v_2, \ldots, v_n\}$ be the row and column vectors of A respectly.

For row operations:

$$
\begin{array}{c}
i-\text{th} \\
\\
j-\text{th}
\end{array}
\begin{pmatrix}
\ddots & & & & \\
& & & 1 & \\
& & \ddots & & \\
& 1 & & & \\
& & & & \ddots
\end{pmatrix}
A =
\begin{pmatrix}
\ddots & & & \\
& - & u_j & - & \\
& & \ddots & & \\
& - & u_i & - & \\
& & & & \ddots
\end{pmatrix}
$$

$$
\begin{array}{c}
\\
i-\text{th} \\
\\
\end{array}
\begin{pmatrix}
\ddots & & & \\
& 1 & & \\
& & c & \\
& & & 1 \\
& & & & \ddots
\end{pmatrix}
A =
\begin{pmatrix}
\ddots & & & \\
& \ddots & & \\
& - & cu_i & - \\
& & & \ddots \\
& & & & \ddots
\end{pmatrix}
$$

$$
\begin{array}{c}
i-\text{th} \\
\\
j-\text{th}
\end{array}
\begin{pmatrix}
\ddots & & & \\
& 1 & & \\
& & \ddots & \\
& c & & 1 \\
& & & & \ddots
\end{pmatrix}
A =
\begin{pmatrix}
\ddots & & & & \\
& \ddots & & & \\
& & \ddots & & \\
& - & cu_i & + & u_j & - \\
& & & & \ddots
\end{pmatrix}
$$

For column operations:

$$
A\begin{pmatrix}
\ddots & & & & \\
 & & & 1 & \\
 & & \ddots & & \\
 & 1 & & & \\
 & & & & \ddots
\end{pmatrix}
\begin{matrix} {}^{i\text{-th}} & {}^{j\text{-th}} \end{matrix}
=
\begin{pmatrix}
\ddots & & & & \\
 & | & & | & \\
 & v_j & \ddots & v_i & \\
 & | & & | & \\
 & & & & \ddots
\end{pmatrix}
$$

$$
A\begin{pmatrix}
\ddots & & & \\
 & 1 & & \\
 & & c & \\
 & & 1 & \\
 & & & \ddots
\end{pmatrix}
\begin{matrix} {}^{i\text{-th}} \end{matrix}
=
\begin{pmatrix}
\ddots & & \\
 & | & \\
 & cv_i & \\
 & | & \ddots \\
 & & & \ddots
\end{pmatrix}
$$

$$
A\begin{pmatrix}
\ddots & & & \\
 & 1 & & c \\
 & & \ddots & \\
 & & & 1 \\
 & & & & \ddots
\end{pmatrix}
\begin{matrix} {}^{i\text{-th}} & {}^{j\text{-th}} \end{matrix}
=
\begin{pmatrix}
\ddots & & & \\
 & & & | \\
 & \ddots & & cv_i \\
 & & \ddots & + \\
 & & & v_j \\
 & & & | & \ddots
\end{pmatrix}
$$

8. By Theorem 3.2 E^{-1} is an elementary matrix of the same type if E is. So if Q can be obtained from P, we can write $Q = EP$ and hence $E^{-1}Q = P$. This means P can be obtained from Q.

9. The operation of interchanging the i-th and the j-th row can be obtained by the following steps:

   - multiplying the i-th row by -1;
   - adding -1 time the i-th row to the j-th row;
   - adding 1 time the j-th row to the i-th row;
   - adding -1 time the i-th row to the j-th row.

10. The operation of multiplying one row by a scalar c means dividing the same row by a scalar $\frac{1}{c}$.

11. The operation of adding c times of the i-th row to the j-th row means substracting c times of the i-th row to the j-th row.

12. Assuming $k = \min\{m,n\}$. Set j be a integer variable and do repeatly the following process:

- If $A_{ij} = 0$ for all j, take $i = i + 1$ and omit following steps and repeat process directly.

- If $A_{ij} \neq 0$ for some j, interchange the i-th and the j-th row.

- Adding $-\frac{A_{ij}}{A_{ii}}$ times the i-th row to the j-th row for all $j > i$.

- Set $i = i + 1$ and repeat the process.

3.2   The Rank of a Matrix and Matrix Inverses

1. (a) No. For example the rank of a 2×2 matrix with all entries 1 has only rank 1.

   (b) No. We have the example that the product of two nonzero matrices could be a zero matrix.

   (c) Yes.

   (d) Yes. This is Theorem 3.4.

   (e) No. They do.

   (f) Yes. This is the Corollary 2 after the Theorem 3.6.

   (g) Yes. This is the argument after the definition of augmented matrix.

   (h) Yes. Rank of an $m \times n$ matrix must be less than m and n.

   (i) Yes. This means L_A is a surjective transformation from $\mathbb{F}^n$ to $\mathbb{F}^n$ and hence a injective transformation, where A is the matrix.

2. In the following questions we may do the Gaussian elimination and the number of nonzero row vectors is equal to the rank of that matrix.

   (a) The rank is 2.
   $$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

   (b) The rank is 3.
   $$\begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

   (c) The rank is 2.
   $$\begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{pmatrix}$$

   (d) The rank is 1.
   $$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

(e) The rank is 3.

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

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

(f) The rank is 3.

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

$$
\rightsquigarrow
\begin{pmatrix}
1 & 2 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 & -2 \\
0 & 0 & 0 & 0 & 2 \\
0 & 0 & 0 & 0 & 7
\end{pmatrix}
\rightsquigarrow
\begin{pmatrix}
1 & 2 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 & -2 \\
0 & 0 & 0 & 0 & 2 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}
$$

(g) The rank is 1.

$$
\begin{pmatrix}
1 & 1 & 0 & 1 \\
2 & 2 & 0 & 2 \\
1 & 1 & 0 & 1 \\
1 & 1 & 0 & 1
\end{pmatrix}
\rightsquigarrow
\begin{pmatrix}
1 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
$$

3. It's natural that $\mathrm{rank}(A) = 0$ if $A = 0$. For the converse, we know that if A is not a zero matrix, we have $A_{ij} \neq 0$ and thus the i-th row is an independent set. So $\mathrm{rank}(A)$ can not be zero.

4. Just do row and column operations.

   (a) The rank is 2.

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

$$
\rightsquigarrow
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
$$

(b) The rank is 2.

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

5. For these problems, we can do the Gaussian elimination on the augment matrix. If the matrix is full rank[1] then we get the inverse matrix in the augmenting part.

(a) The rank is 2 and its inverse is $\begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$.

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -1 & -1 & 1 \end{array} \right)$$

$$\rightsquigarrow \left(\begin{array}{cc|cc} 1 & 0 & -1 & 2 \\ 0 & -1 & -1 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 0 & -1 & 2 \\ 0 & 1 & 1 & -1 \end{array} \right)$$

(b) The rank is 1. So there's no inverse matrix.

(c) The rank is 2. So there's no inverse matrix.

(d) The rank is 3 and its inverse is $\begin{pmatrix} -\frac{1}{2} & 3 & -1 \\ \frac{3}{2} & -4 & 2 \\ 1 & -2 & 1 \end{pmatrix}$.

(e) The rank is 3 and its inverse is $\begin{pmatrix} \frac{1}{6} & -\frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{6} & \frac{1}{3} & \frac{1}{2} \end{pmatrix}$.

(f) The rank is 2. So there's no inverse matrix.

(g) The rank is 4 and its inverse is $\begin{pmatrix} -51 & 15 & 7 & 12 \\ 31 & -9 & -4 & -7 \\ -10 & 3 & 1 & 2 \\ -3 & 1 & 1 & 1 \end{pmatrix}$.

(h) The rank is 3. So there's no inverse matrix.

6. For these problems we can write down the matrix representation of the transformation $[T]_\alpha^\beta$, where $\alpha = \{u_1, u_2, \ldots, u_n\}$ and $\beta = \{v_1, v_2, \ldots, v_n\}$ are the (standard) basis for the domain and the codomain of T. And the inverse of this matrix would be $B = [T^{-1}]_\beta^\alpha$. So T^{-1} would be the linear transformation such that

$$T^{-1}(v_j) = \sum_{i=1}^n B_{ij} u_i.$$

---

[1]This means $\text{rank}(A) = n$, where A is an $n \times n$ matrix.

(a) We get $[T]_\alpha^\beta = \begin{pmatrix} -1 & 2 & 2 \\ 0 & -1 & 4 \\ 0 & 0 & -1 \end{pmatrix}$ and $[T^{-1}]_\beta^\alpha = \begin{pmatrix} -1 & -2 & -10 \\ 0 & -1 & -4 \\ 0 & 0 & -1 \end{pmatrix}$. So we know that

$$T(a + bx + cx^2) = (-a - 2b - 10c) + (-b - 4c)x + (-c)x^2.$$

(b) We get $[T]_\alpha^\beta = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}$, a matrix not invertible. So T is not invertible.

(c) We get $[T]_\alpha^\beta = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}$ and $[T^{-1}]_\beta^\alpha = \begin{pmatrix} \frac{1}{6} & -\frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{6} & \frac{1}{3} & \frac{1}{2} \end{pmatrix}$. So we know that

$$T(a, b, c) = (\frac{1}{6}a - \frac{1}{3}b + \frac{1}{2}c, \frac{1}{2}a - \frac{1}{2}c, -\frac{1}{6}a + \frac{1}{3}b + \frac{1}{2}c).$$

(d) We get $[T]_\alpha^\beta = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ and $[T^{-1}]_\beta^\alpha = \begin{pmatrix} 0 & 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & -1 \end{pmatrix}$. So we know that

$$T(a, b, c) = (c, \frac{1}{2}a - \frac{1}{2}b, \frac{1}{2}a + \frac{1}{2}b - c).$$

(e) We get $[T]_\alpha^\beta = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$ and $[T^{-1}]_\beta^\alpha = \begin{pmatrix} 0 & 1 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & -1 & \frac{1}{2} \end{pmatrix}$. So we know that

$$T(a + bx + cx^2) = (b, -\frac{1}{2}a + \frac{1}{2}c, \frac{1}{2}a - b + \frac{1}{2}c).$$

(f) We get $[T]_\alpha^\beta = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$, a matrix not invertible. So T is not invertible.

7. We can do the Gaussian elimination and record what operation we've done.

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

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

$$\rightsquigarrow \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Let $E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$, $E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $E_4 = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$,

$E_5 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$, $E_6 = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Thus we have the matrix equals to $E_1^{-1} E_2^{-1} E_3^{-1} E_4^{-1} E_5^{-1} E_6^{-1}$.

8. It's enough to show that $R(L_A) = R(L_{cA})$. But this is easy since

$$R(L_A) = L_A(\mathbb{F}^m) = cL_A(\mathbb{F}^m) = L_{cA}(\mathbb{F}^m) = R(L_cA).$$

9. If B is obtained from a matrix A by an elementary column operation, then there exists an elementary matrix such that $B = AE$. By Theorem 3.2, E is invertible, and hence rank(B) =rank(A) by Theorem 3.4.

10. Let A be an $m \times 1$ matrix. Let i be the smallest integer such that $A_{i1} \neq 0$. Now we can interchange, if it's necessary, the first and the i-th row. Next we can multiply the first row by scalar $\frac{1}{A_{11}}$ and we get $A_{11} = 1$ now. Finally add $-A_{k1}$ times the first row to the k-th row. This finished the process.

11. We may write $B_0' = \begin{pmatrix} 0 & \cdots & 0 \\ & B' & \end{pmatrix}$. And thus we have that

$$B \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n+1} \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + B_0' \begin{pmatrix} x_2 \\ \vdots \\ x_{n+1} \end{pmatrix}.$$

Let L =span$\left\{\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\right\}$ be a subspace in $\mathbb{F}^{m+1}$. So we have that $R(L_B) = L + R(B_0')$. And it's easy to observe that all element in L has its first entry nonzero except 0. But all element in $R(B_0')$ has it first entry zero. So we know that $L \cap R(B_0') = \{0\}$ and hence $R(L_B) = L \oplus R(B_0')$. By Exercise 1.6.29(b) we know that dim$(R(L_B))$ =dim(L)+dim$(R(B_0'))$ = 1+dim$(R(B_0'))$.

Next we want to prove that $\dim(R(B'_0)) = \dim(B')$ by showing $N(L_{B'_0}) = N(L_{B'})$. We may let

$$B'_0 x = \begin{pmatrix} 0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}$$

and scrutinize the fact

$$B' x = \begin{pmatrix} 0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}$$

is true. So $N(L_{B'_0}) = N(L_{B'})$ is an easy result of above equalitiies. Finally since $L_{B'_0}$ and L'_B has the same domain, by Dimension Theorem we get the desired conclusion

$$\operatorname{rank}(B) = \dim(R(L_B)) = 1 + \dim(R(B'_0))$$

$$= 1 + \dim(L_{B'}) = 1 + \operatorname{rank}(B').$$

12. If B' can be transformed into D' by an elementary row operation, we could write $D' = EB'$ by some elementary matrix E. Let

$$E' = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B' & \\ 0 & & & \end{array} \right)$$

be an larger matrix. Then we have $D = E'B$ and hence D can be obtained from B by an elementary row operation. For the version of column, we have $D' = B'E$. And then get the matrix E' by the same way. Finally we have $D = BE'$.

13. **(b)** By Theorem 3.5 and the Corollary 2 after Theorem 3.6 we have that the maximum number of linearly independent rows of A is the maximum number of linearly independent columns of A^t and hence equals to $\operatorname{rank}(A^t) = \operatorname{rank}(A)$.

    (c) This is an instant result of (b) and Theorem 3.5.

14. (a) For all $y \in R(T + U)$ we can express it as $y = T(x) + U(x) \in R(T) + R(U)$ for some $x \in V$.

    (b) By Theorem 1.6.29(a) we have

$$\operatorname{rank}(T + U) \le \dim(R(T) + R(U))$$

$$= \dim(R(T)) + \dim(R(U)) - \dim(R(T) \cap R(U))$$

$$\le \operatorname{rank}(T) + \operatorname{rank}.$$

74

(c) We have that

$$\text{rank}(A + B) = \text{rank}(L_{A+B}) = \text{rank}(L_A + L_B)$$

$$\le \text{rank}(A) + \text{rank}(B).$$

15. Let $P = M(A|B)$ and $Q = (MA|MB)$. We want to show that $P_{ij} = Q_{ij}$ for all i and for all j. Assume that A and B has a and b columns respectly. For $j = 1, 2, \ldots, a$, we have that

$$P_{ij} = \sum_{k=1}^{n} M_{ik} A_{kj} = (MA)_{ij} = Q_{ij}.$$

For $j = a + 1, a + 2, \ldots, a + b$, we have that

$$P_{ij} = \sum_{k=1}^{n} M_{ik} B_{kj} = (MB)_{ij} = Q_{ij}.$$

16. Since P is invertible, we know that L_P is an isomorphism. So by Exercise 2.4.17 we have that

$$\text{rank}(PA) = \dim(P(A(\mathbb{F}^n)))$$

$$= \dim(A(\mathbb{F}^n)) = A(\mathbb{F}^n)(A).$$

17. Let $B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ and $C = \begin{pmatrix} c_1 & c_2 & c_3 \end{pmatrix}$. Thus we know that

$$BC = \begin{pmatrix} b_1 c_1 & b_1 c_2 & b_1 c_3 \\ b_2 c_1 & b_2 c_2 & b_2 c_3 \\ b_3 c_1 & b_3 c_2 & b_3 c_3 \end{pmatrix}$$

has only at most one independent rows. So the rank of BC is at most one.

Conversely, if the rank of A is zero, we know that $A = O$ and we can pick B and C such that they are all zero matrices. So assume that the rank of A is 1 and we have that the i-th row of A forms an maximal independent set itself. This means that we can obtained the other row of A by multiplying some scalar (including 0), say b_j for the j-the row. Then we can pick $B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ and $C = \begin{pmatrix} A_{i1} \\ A_{i2} \\ A_{i3} \end{pmatrix}$. Thus we get the desired matrices.

18. Let A_i be the matrix consists of the i-th column of A. Let B_i be the matrix consists of the i-th row of B. It can be check that actually

$$AB = \sum_{i=1}^{n} A_i B_i$$

and $A_i B_i$ has rank at most 1.

19. It would be m. Since the range of A is a subspace with dimension m in $\mathbb{F}^m$, we know that $L_A(\mathbb{F}^n) = \mathbb{F}^>$. Similarly we also have that $L_B(\mathbb{F}^p) = \mathbb{F}^n$. So we know that

$$L_AB(\mathbb{F}^p) = L_A(L_B(\mathbb{F}^p)) = \mathbb{F}^m$$

and hence the rank of AB is m.

20. (a) Just like the skill we learned in the Exercise 1.4.2 we can solve the system of linear equation $Ax = 0$ and get the solution space

$$\{(x_3 + 3x_5, -2x_3 + x_5, x_3, -2x_5, x_5) : x_i \in \mathbb{F}\}.$$

And we know that $\{(1, -2, 1, 0, 0), (3, 1, 0, -2, 1)\}$ is a basis for the solution space. Now we can construct the desired matrix

$$M = \begin{pmatrix} 1 & 3 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

(b) If $AB = O$, this means that every column vector of B is a solution of $Ax = 0$. If rank of B is greater than 2, we can find at least three independent vectors from columns of B. But this is is impossible since by Dimension Theorem we know that

$$\dim(\mathbb{F}^5) = \dim(R(L_A)) + \dim(N(L_A))$$

and so $\dim(N(L_A)) = 5 - 3 = 2$.

21. Let $\beta = \{e_1, e_2, \ldots, e_m\}$ be the standard basis for $\mathbb{F}^m$. Since the rank of A is m, we know that L_A is surjective. So we can find some vector $v_i \in \mathbb{F}^n$ such that $L_A(v_i) = e_i$. So let B be the matrix with column vector v_i. Thus B is an $n \times m$ matrix and $AB = I$ since $Av_i = e_i$.

22. We know that B^t is an $m \times n$ matrix with rank n. By the previous exercise we have some $n \times m$ matrix C such that $B^t C = I_m$. We may pick $A = C^t$. Now we have the fact that $AB = C^t B = (B^t C)^t = (I_m)^t - I_m$.

3.3 Systems of Linear Equation—Theoretical Aspects

1. (a) No. The system that $0x = 1$ has no solution.

   (b) No. The system that $0x = 0$ has lots of solutions.

   (c) Yes. It has the zero solution.

   (d) No. The system that $0x = 0$ has no solution.

(e) No. The system that $0x = 0$ has lots of solutions.

(f) No. The system $0x = 1$ has no solution but the homogeneous system corresponding to it has lots of solution.

(g) Yes. If $Ax = 0$ then we know $x = A^{-1}0 = 0$.

(h) No. The system $x = 1$ has solution set $\{1\}$.

2. See Example 2 in this section.

(a) The set $\{(-3,1)\}$ is a basis and the dimension is 1.

(b) The set $\{(\frac{1}{3}, \frac{2}{3}, 1)\}$ is a basis and the dimension is 1.

(c) The set $\{(-1,1,1)\}$ is a basis and the dimension is 1.

(d) The set $\{(0,1,1)\}$ is a basis and the dimension is 1.

(e) The set $\{(-2,1,0,0), (3,0,1,0), (-1,0,0,1)\}$ is a basis and the dimension is 3.

(f) The set $\{(0,0)\}$ is a basis and the dimension is 0.

(g) The set $\{(-3,1,1,0), (1,-1,0,1)\}$ is a basis and the dimension is 1.

3. See Example 3 in this section.

(a) The solution set is $(5,0) + \text{span}(\{(-3,1)\})$.

(b) The solution set is $(\frac{2}{3}, \frac{1}{3}, 0) + \text{span}(\{(\frac{1}{3}, \frac{2}{3}, 1)\})$.

(c) The solution set is $(3,0,0) + \text{span}(\{(-1,1,1)\})$.

(d) The solution set is $(2,1,0) + \text{span}(\{(0,1,1)\})$.

(e) The solution set is $(1,0,0,0) + \text{span}(\{(-2,1,0,0), (3,0,1,0), (-1,0,0,1)\})$.

(f) The solution set is $(1,2) + \text{span}(\{(0,0)\})$.

(g) The solution set is $(-1,1,0,0) + \text{span}(\{(-3,1,1,0), (1,-1,0,1)\})$.

4. With the technique used before we can calculate A^{-1} first and then the solution of $Ax = b$ would be $A^{-1}b$ if A is invertible.

(a) Calculate $A^{-1} = \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix}$ and solution is $x_1 = -11$, $x_2 = 5$.

(b) Calculate $A^{-1} = \begin{pmatrix} \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{9} & \frac{1}{3} & -\frac{2}{9} \\ -\frac{4}{9} & \frac{2}{3} & -\frac{1}{9} \end{pmatrix}$ and solution is $x_1 = 3$, $x_2 = 0$, $x_3 = -2$.

5. Let A be the $n \times n$ zero matrix. The system $Ax = 0$ has infinitely many solutions.

6. If $T(a,b,c) = (a+b, 2a-c) = (1,11)$, then we get $a + b = 1$ and $2a - c = 11$. This means the preimage set would be

$$T^{-1}(1,11) = \{(a, 1-a, 2a-11) : a \in \mathbb{R}\}.$$

7. See Theorem 3.11 and Example 5 of this section.

    (a) It has no solution.

    (b) It has a solution.

    (c) It has a solution.

    (d) It has a solution.

    (e) It has no solution.

8. (a) Just solve that $a + b = 1$, $b - 2c = 3$, $a + 2c = -2$ and get $a = 0$, $b = 1$, $c = -1$ is a solutoin. So we know $v \in R(T)$.

    (b) Just solve that $a + b = 2$, $b - 2c = 1$, $a + 2c = 1$ and get $a = 1$, $b = 1$, $c = 0$ is a solution. So we know $v \in R(T)$.

9. This is the definition of L_A and $R(L_A)$.

10. The answer is Yes. Say the matrix is A. Since the matrix has rank m, we have that dimension of $R(L_A)$ is m. But this means $R(L_A) = \mathbb{F}^m$ since the codomain of L_A is $\mathbb{F}^m$ and it has dimension m. So it must has a solution by the previous exercise.

11. Solve the system $Ax = x$ and we can get $x = (\frac{4}{11}, \frac{3}{11}, \frac{4}{11})$. And the amount of each entry is the ratio of farmer, trailor, and carpenter respectly.

12. Set
$$A = \begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix}$$

be the input-output matrix of this system. We want to solve that $Ax = x$, which means $(A - I)x = 0$. By calculation we get $x = t(3, 4)$ for arbitrary $t \in \mathbb{R}$. So the proportion is used in the production of goods would be $\frac{3}{7}$.

13. In this model we should solve the equation $(I - A)x = d$. And we can compute that $I - A$ is invertible and
$$(I - A)^{-1} = \begin{pmatrix} \frac{12}{5} & \frac{3}{5} \\ 1 & \frac{3}{2} \end{pmatrix}.$$

So we know
$$x = (I - A)^{-1}d = \begin{pmatrix} \frac{39}{5} \\ \frac{19}{2} \end{pmatrix}.$$

14. The input-output matrix A should be $\begin{pmatrix} 0.50 & 0.20 \\ 0.30 & 0.60 \end{pmatrix}$. And the demand vector d should be $\begin{pmatrix} 90 \\ 20 \end{pmatrix}$. So we can solve the equation $(I - A)x = d$ and get the answer $x = (\frac{2000}{7}, \frac{1850}{7})$. Thus x is the support vector.

3.4 Systems of Linear Equations—Computational Aspects

1. (a) No. This form could only fit for row operation. For example, $x = 1$ and $x = 0$ has different solution set.

    (b) Yes. This is Theorem 3.13.

    (c) Yes. This is the result of Theorem 3.16.

    (d) Yes. This is Theorem 3.14.

    (e) No. For example, the system with corresponding augmented matrix $\begin{pmatrix} 0 & | & 1 \end{pmatrix}$ has no solution.

    (f) Yes. This is Theorem 3.15.

    (g) Yes. This is Theorem 3.16.

2. (a) The solution is $(4, -3, , -1)$.

    (b) The solution set is $\{(9, 4, 0) + t(-5, -3, 1)\}$.

    (c) The solution is $(2, 3, -2, 1)$.

    (d) The solution is $(-21, -16, 14, -10)$.

    (e) The solution set is $\{(4, 0, 1, 0) + s(4, 1, 0, 0) + t(1, 0, 2, 1)\}$.

    (f) The solution set is $\{(-3, 3, 1, 0) + t(1, -2, 0, 1)\}$.

    (g) The solution set is $\{(-23, 0, 7, 9, 0) + s(0, 2, 1, 0, 0) + t(-23, 0, 6, 9, 1)\}$.

    (h) The solution set is $\{(-3, -8, 0, 0, 3) + t(1, -2, 1, 0, 0)\}$.

    (i) The solution set is $\{(2, 0, 0, -1, 0) + s(0, 2, 1, 0, 0) + t(1, -1, 1, 2)\}$.

    (j) The solution set is $\{(1, 0, 1, 0) + s(-1, 1, 0, 0) + t(1, 0, 1, 2)\}$.

3. (a) We can check that A' is also reduced echelon form. So the number of nonzero rows in A' is the rank of A. And the number of nonzero rows in $(A'|b')$ is the rank of $(A|b)$. So if they have different rank there must contain some nonzero rows (actually only one row) in $(A'|b')$ but not in A'. This means the nonzero row must has nonzero entry in the last column. Conversely, if some row has its only nonzero entry in the last column, this row did not attribute the rank of A'. Since every nonzero row in A' has its corresponding row in $(A'|b')$ also a nonzero row, we know that two matrix have different rank.

    (b) By the previous exercise we know that $(A'|b')$ contains a row with only nonzero entry in the last column is equivalent to that A' and $(A'|b')$ have different rank. With the help of Theorem 3.11 we get the desired conclusion.

4. (a) The solution set is $\{(\frac{4}{3}, \frac{1}{3}, 0, 0) + t(1, -1, 1, 2)\}$. The basis is $\{(1, -1, 1, 2)\}$.

    (b) The solution set is $\{(1, 0, 1, 0) + s(-1, 1, 0, 0) + t(1, 0, 1, 2)\}$. The basis is $\{(-1, 1, 0, 0), (1, 0, 1, 2)\}$.

(c) It has no solution.

5. Let R be the matrix in reduced echelon form. We know that there is an invertible matrix C such that $CA = R$. This means

$$C \begin{pmatrix} 1 & 0 & 1 \\ -1 & -1 & -2 \\ 3 & 1 & 0 \end{pmatrix} = I.$$

So we get

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

And hence

$$A = C^{-1}R = \begin{pmatrix} 1 & 0 & 2 & 1 & 4 \\ -1 & -1 & 3 & -2 & -7 \\ 3 & 1 & 1 & 0 & -9 \end{pmatrix}.$$

6. Let R be the matrix in reduced echelon form. We know that there is an invertible matrix C such that $CA = R$. But now we cannot determine what C is by the given conditions. However we know that the second column of R is -3 times the first column of R. This means

$$0 = R \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = CA \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Since C is invertible, we know that

$$A \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 0.$$

And this means the second column of A is also -3 times the first column of A. And so the second column of A is $(-3, 6, 3, -9)$. Similarly we have that

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

and get the answer that matrix A is

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

7. See Exercise 1.6.8. Note that if we put those vector as row vectors of M just like what we've done in the Exercise 1.6.8, we cannot interchange any two rows. However we can also construct a matrix the i-th column u_i. And we can do the row operation including interchanging any two rows. The set of columns containing one pivot[2] forms an independent set.

$$\begin{pmatrix} 2 & 1 & -8 & 1 & -3 \\ -3 & 4 & 12 & 37 & -5 \\ 1 & -2 & -4 & -17 & 8 \end{pmatrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & \frac{1}{2} & -4 & \frac{1}{2} & -\frac{3}{2} \\ 0 & 1 & 0 & 7 & -\frac{19}{11} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

So the set $\{u_1, u_2, u_5\}$ is a basis for $\mathbb{R}^3$.

8. Do the same just like what we've done in the previous exercise.

$$\begin{pmatrix} 2 & -6 & 3 & 2 & -1 & 0 & 1 & 2 \\ -3 & 9 & -2 & -8 & 1 & -3 & 0 & -1 \\ 4 & -12 & 7 & 2 & 2 & -18 & -2 & 1 \\ -5 & 15 & -9 & -2 & 1 & 9 & 3 & -9 \\ 2 & -6 & 1 & 6 & -3 & 12 & -2 & 7 \end{pmatrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & -3 & \frac{3}{2} & 1 & -\frac{1}{2} & 0 & \frac{1}{2} & 1 \\ 0 & 0 & 1 & -2 & -\frac{1}{5} & -\frac{6}{5} & \frac{3}{5} & \frac{4}{5} \\ 0 & 0 & 0 & 0 & 1 & -4 & -\frac{23}{21} & -\frac{19}{21} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

We know that $\{u_1, u_3, u_5, u_7\}$ is a basis for W.

9. Use the representation of those matrix with some basis (usually take the standard basis) and do the same thing like previous questions on them.

$$\begin{pmatrix} 0 & 1 & 2 & 1 & -1 \\ -1 & 2 & 1 & -2 & 2 \\ -1 & 2 & 1 & -2 & 2 \\ 1 & 3 & 9 & 4 & -1 \end{pmatrix}$$

---

[2]The position who is the first nonzero entry in one nonzero row is called a pivot. For example, the position 11, 22, 35 are pivots.

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

So we know that the subset containing the first, the second, and the fourth matrix forms a basis for W.

10. (a) It's easy to check that

$$0 - 2 + 3 - 1 - 0 + 0 = 0.$$

So the vector $(0, 1, 1, 1, 0)$ is an element in V. Since the set contains only one nonzero vector, it's linearly independent.

(b) As usual we can find a basis

$$\beta = \{(2, 1, 0, 0, 0), (-3, 0, 1, 0, 0), (1, 0, 0, 1, 0), (-2, 0, 0, 0, 1)\}.$$

So we know that $\{(0, 1, 1, 1, 0)\} \cup \beta$ can generate the space V. Do the same thing to this new set and remember to put $(0, 1, 1, 1, 0)$ on the first column in order to keep it as an element when we do Gaussian elimination.

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

$$\rightsquigarrow \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Now we know that

$$\beta' = \{(0, 1, 1, 1, 0), (2, 1, 0, 0, 0), (-3, 0, 1, 0, 0), (-2, 0, 0, 0, 1)\}$$

forms a basis for V.

11. (a) Similarly check

$$1 - 4 + 3 + 0 - 0 = 0.$$

So the set containing only the vector $(1, 2, 1, 0, 0)$ is linearly independent by the same reason.

(b) Do the same thing to the set $\{(1, 2, 1, 0, 0)\} \cup \beta$.

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

$$\leadsto \begin{pmatrix} 1 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Now we know that the set

$$\{(1,2,1,0,0), (2,1,0,0,0), (1,0,0,1,0), (-2,0,0,0,1)\}$$

forms a basis for V.

12. (a) Set $v_1 = (0,-1,0,1,1,0)$ and $v_2 = (1,0,1,1,1,0)$. Check the two vectors satisfy the system of linear equation and so they are vectors in V. To show they are linearly independent, assume that

$$a(0,-1,0,1,1,0) + b(1,0,1,1,1,0)$$

$$= (b, -a, b, a+b, a+b, 0) = 0.$$

This means that $a = b = 0$ and the set is independent.

(b) Similarly we find a basis

$$\beta = \{(1,1,1,0,0,0), (-1,1,0,1,0,0)$$

$$, (1,-2,0,0,1,0), (-3,-2,0,0,0,1)\}$$

for V as what we do in the Exercise 3.4.4. Still remember that we should put v_1 and v_2 on the first and the second column.

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

$$\leadsto \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

So the set

$$\{(0,-1,0,1,1,0), (1,0,1,1,1,0), (-1,1,0,1,0,0), (-3,-2,0,0,0,1)\}$$

forms a basis for V.

13. (a) Set $v_1 = (1, 0, 1, 1, 1, 0)$ and $v_2 = (0, 2, 1, 1, 0, 0)$. Check the two vectors satisfy the system of linear equation and so they are vectors in V. To show they are linearly independent, assume that

$$a(1, 0, 1, 1, 1, 0) + b(0, 2, 1, 1, 0, 0)$$

$$= (a, 2b, a + b, a + b, a, 0) = 0.$$

This means that $a = b = 0$ and the set is independent.

(b) Take the same basis β as that in the previous exercise and do Gaussian elimination.

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

$$\rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

So the set

$$\{(1, 0, 1, 1, 1, 0), (0, 2, 1, 1, 0, 0), (1, 1, 1, 0, 0, 0), (-3, -2, 0, 0, 0, 1)\}$$

forms a basis for V.

14. It's enough to check that A satisfies the definition of reduced echelon form on the page 185. For the first condition, a nonzero row in A is a nonzero row in $(A|b)$. So it will precede all the zero rows in $(A|b)$. But there may be some zero rows in A who are nonzero rows in $(A|b)$. This kind of rows will be behind those nonzero row by the third condition for $(A|b)$. The second condition for $(A|b)$ implies the second condition for A. The third condition for $(A|b)$ implies the third condition for A.

15. We call a column whose corresponding column in one fixed reduced echelon form contains a pivot(see Exercise 3.4.7) a pivotal column. Now we induct on the number of columns of a matrix. For matrix contains only one column u_1, the reduced echelon form of it would be the column e_1 if u_1 is nonzero (and hence $\{u_1\}$ is independent) and that of it would be the zero column if u_1 is zero (and hence $\{u_1\}$ is dependent). Suppose that the reduced echelon form of a matrix with k columns is unique. Now consider a matrix A with $k + 1$ columns, say $u_1, u_2, \ldots, u_{k+1}$. Let A' be the matrix by deleting the final column of A. So we can write $A = (A'|u_{k+1})$. Say $(R'|b)$ is a reduced echelon form of A. By the previous exercise we

know that R' is a reduced echelon form of A'. And R' is unique by our induction hypothesis. So the set of pivotal columns P' in A' is also unique. By Theorem 3.16(c) and Theorem 3.16(d) we know that the set P' in A' is a maximal independent set of $\{u_1, u_2, \ldots, u_k\}$. Now if $P' \cup \{u_k\}$ is linearly independent, this means

$$\mathrm{rank}(A) = \mathrm{rank}(A') + 1,$$

say the value is r. By Theorem 3.16(a) and Theorem 3.16(b), b is the only column who can be e_r and so we know that $b = e_r$. On the other hand, if $P' \cup \{u_k\}$ linearly dependent. The vector u_k cannot be a pivotal column of A since $P' \cup \{u_k\}$ is the set of pivotal column of A and it must be linearly independent. Futhermore, u_k has an unique representation with respect to the set P'. By Theorem 3.16(d), the column vector d must be the representation of u_k. By all cases we know that $(R'|b)$ is also unique. And by induction we get the desired conclusion.

Chapter 4

Determinants

4.1 Determinants of Order 2

1. (a) No. We have $\det(2I_2) = 4$ but not $2\det(T_2)$.

   (b) Yes. Check that

   $$\det\begin{pmatrix} a_1 + ka_2 & b_1 + kb_2 \\ c & d \end{pmatrix} = (a_1 d - b_1 c) + k(a_2 d - b_2 c)$$

   $$= \det\begin{pmatrix} a_1 & b_1 \\ c & d \end{pmatrix} + k\det\begin{pmatrix} a_2 & b_2 \\ c & d \end{pmatrix}$$

   and

   $$\det\begin{pmatrix} a & b \\ c_1 + kc_2 & d_1 + kd_2 \end{pmatrix} = (ad_1 - bc_1) + k(ad_2 - bc_2)$$

   $$= \det\begin{pmatrix} a & b \\ c_1 & d_1 \end{pmatrix} + k\det\begin{pmatrix} a & b \\ c_2 & d_2 \end{pmatrix}$$

   for every scalar k.

   (c) No. A is invertible if and only if $\det(A) \neq 0$.

   (d) No. The value of the area cannot be negative but the value $\det\begin{pmatrix} u \\ v \end{pmatrix}$ could be.

   (e) Yes. See Exercise 4.1.12.

2. Use the formula in Definition in page 200.

   (a) The determinant is $6 \times 4 - (-3) \times 2 = 30$.

   (b) The determinant is -17.

   (c) The determinant is -8.

3. (a) The determinant is $-10 + 15i$.

86

(b) The determinant is $-8 + 29i$.

(c) The determinant is -24.

4. Compute $\left| \det \begin{pmatrix} u \\ v \end{pmatrix} \right|$.

(a) The area is $|3 \times 5 - (-2) \times 2| = 19$.

(b) The area is 10.

(c) The area is 14.

(d) The area is 26.

5. It's directly from the fact
$$\det \begin{pmatrix} c & d \\ a & b \end{pmatrix} = cb - da = -(ad - bc) = -\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

6. It's directly from the fact
$$\det \begin{pmatrix} a & b \\ a & b \end{pmatrix} = ab - ba = 0.$$

7. It's directly from the fact
$$\det \begin{pmatrix} a & c \\ b & d \end{pmatrix} = ad - cb = ad - bc = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

8. It's directly from the fact
$$\det \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = ad - b0 = ad.$$

9. Directly check that
$$\det\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} \right) = \det \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix}$$
$$= (ae+bg)(cf+dh)-(af+bh)(ce+dg) = ad(eh-fg)-bc(eh-fg) = (ad-bc)(eh-fg)$$
$$= \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \times \det \begin{pmatrix} e & f \\ g & h \end{pmatrix}.$$

10. For brevity, we write $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ for some 2×2 matrix and $C = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$ for the corresponding classical adjoint.

(a) Directly check that

$$CA = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix}$$

and

$$AC = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix}.$$

(b) Calculate that

$$\det(C) = da - (-c)(-b) = ad - bc = \det(A).$$

(c) Since the transpose matrix A^t is $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$, the corresponding classical adjoint would be

$$\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = C^t.$$

(d) If A is invertible, we have that $\det(A) \neq 0$ by Theorem 4.2. So we can write

$$[\det(A)]^{-1}CA = A[\det(A)]^{-1}C = I$$

and get the desired result.

11. By property (ii) we have the fact

$$\delta \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = 0 = \delta \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

Since by property (i) and (ii) we know

$$0 = \delta \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \delta \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} + \delta \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

$$= \delta \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + \delta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \delta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \delta \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$

$$\delta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \delta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

we get that

$$\delta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = -\delta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = -1$$

by property (iii). Finally, by property (i) and (iii) we can deduce the general formula for δ below.

$$\delta \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a\delta \begin{pmatrix} 1 & 0 \\ c & d \end{pmatrix} + b\delta \begin{pmatrix} 0 & 1 \\ c & d \end{pmatrix}$$

$$= ac\delta \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + ad\delta \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} + bc\delta \begin{pmatrix} 0 & 1 \\ c & d \end{pmatrix} + bd\delta \begin{pmatrix} 0 & 1 \\ c & d \end{pmatrix}$$

$$= ad - bc = \det \delta \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

12. A coordinate system $\{u = (a, b), v = (c, d)\}$ is right-handed means $u' \cdot v > 0$ where the vector $u' = (-b, a)$ is obtained by rotating the vector u in a counterclockwise direction through an angle $\frac{\pi}{2}$. With the fact $u' \cdot v = ad - bc = \det \begin{pmatrix} u \\ v \end{pmatrix}$ we get the conclusion.

4.2 Determinants of Order n

1. (a) No. See Exercise 4.1.1(a).

   (b) Yes. This is Theorem 4.4.

   (c) Yes. This is the Corollary after Theorem 4.4.

   (d) Yes. This is Theorem 4.5.

   (e) No. For example, the determinant of $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ is 2 but not $\det(I) = 1$.

   (f) No. We have that $\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = 1 \neq 2 \det(I) = 2$.

   (g) No. For example, the determinant of identity matrix is 1.

   (h) Yes. See Exercise 4.2.23.

2. Determinant is linear when we fixed all but one row. So we have that

$$\det \begin{pmatrix} 3a_1 & 3a_2 & 3a_3 \\ 3b_1 & 3b_2 & 3b_3 \\ 3c_1 & 3c_2 & 3c_3 \end{pmatrix} = 3 \det \begin{pmatrix} a_1 & a_2 & a_3 \\ 3b_1 & 3b_2 & 3b_3 \\ 3c_1 & 3c_2 & 3c_3 \end{pmatrix}$$

$$= 9 \det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ 3c_1 & 3c_2 & 3c_3 \end{pmatrix} = 27 \det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}.$$

   Hene we know that $k = 27$.

3. We can add $-\frac{5}{7}$ times of the third row the the second row without changing the value of determinant and do the same as the previous exercise and get the conclusion that $k = 2 \times 3 \times 7 = 42$.

4. See the following process.

$$\det \begin{pmatrix} b_1 + c_1 & b_2 + c_2 & b_3 + c_3 \\ a_1 + c_1 & a_2 + c_2 & a_3 + c_3 \\ a_1 + b_1 & a_2 + b_2 & a_3 + b_3 \end{pmatrix} = -\det \begin{pmatrix} -(b_1 + c_1) & -(b_2 + c_2) & -(b_3 + c_3) \\ a_1 + c_1 & a_2 + c_2 & a_3 + c_3 \\ a_1 + b_1 & a_2 + b_2 & a_3 + b_3 \end{pmatrix}$$

$$= -2\det\begin{pmatrix} a_1 & a_2 & a_3 \\ a_1+c_1 & a_2+c_2 & a_3+c_3 \\ a_1+b_1 & a_2+b_2 & a_3+b_3 \end{pmatrix} = -2\det\begin{pmatrix} a_1 & a_2 & a_3 \\ c_1 & c_2 & c_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

$$= 2\det\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

The first equality comes from adding one time the second row and one time the third row to the first column and the second equality comes from adding -1 time the first row to the second and the third row. Finally we interchange the second and the third row and multiply the determinant by -1. Hence k would be 2.

5. The determinant should be -12 by following processes.

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

$$= 0\det\begin{pmatrix} 0 & -3 \\ 3 & 0 \end{pmatrix} - 1\det\begin{pmatrix} -1 & -3 \\ 2 & 0 \end{pmatrix} + 2\det\begin{pmatrix} -1 & 0 \\ 2 & 3 \end{pmatrix}$$

$$= 0\times 9 - 1\times 6 + 2\times(-3) = -12$$

6. The determinant should be -13.

7. The determinant should be -12.

8. The determinant should be -13.

9. The determinant should be 22.

10. The determinant should be $4+2i$.

11. The determinant should be -3.

12. The determinant should be 154.

13. The determinant should be -8.

14. The determinant should be -168.

15. The determinant should be 0.

16. The determinant should be 36.

17. The determinant should be -49.

18. The determinant should be 10.

19. The determinant should be $-28-i$.

90

20. The determinant should be $17 - 3i$.

21. The determinant should be 95.

22. The determinant should be -100.

23. Use induction on n, the size of the matrix. For $n = 1$, every 1×1 matrix is upper triangular and we have the fact $\det(a) = a$. Assuming the statement of this exercise holds for $n = k$, consider any $(n+1) \times (n+1)$ upper triangular matrix A. We can expand A along the first row with the formula

$$\det(A) = \sum_{j=1}^{n+1} (-1)^{1+j} A_{1j} \det(\tilde{A}_{1j}).$$

And the matrix $\tilde{A}_{1j}$, $j \neq 1$, contains one zero collumn and hence has rank less than $n + 1$. By the Corollary after Theorem 4.6 those matrix has determinant 0. However, we have the matrix $\tilde{A}_{11}$ is upper triangular and by induction hypothesis we have

$$\det(\tilde{A}_{11}) = \prod_{i=2}^{n+1} A_{ii}.$$

So we know the original formula would be

$$\det(A) = A_{11} \det(\tilde{A}_{11}) = \prod_{i=1}^{n+1} A_{ii}.$$

24. Let z be the zero row vector. Thus we have that

$$\det \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{r-1} \\ z \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{r-1} \\ 0z \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix} = 0 \det \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{r-1} \\ z \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix} = 0$$

25. Applies Theorem 4.3 to each row. Thus we have hat

$$\det \begin{pmatrix} ka_1 \\ ka_2 \\ \vdots \\ ka_n \end{pmatrix} = k \det \begin{pmatrix} a_1 \\ ka_2 \\ \vdots \\ ka_n \end{pmatrix} = \cdots = k^n \det \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

26. By the previous exercise the equality holds only when n is even or $\mathbb{F}$ has characteristic 2.

91

27. If A has two identical columns, then the matrix A is not full-rank. By Corollary after Theorem 4.6 we know the determinant of A should be 0.

28. The matrix E_1 can be obtained from I by interchanging two rows. So by Theorem 4.5 the determinant should be -1. The matrix E_2 is upper triangular. By Exercise 4.2.23 the determinant should be c, the scalar by whom some row was multiplied. The matrix E_3 has determinant 1 by Theorem 4.6.

29. The elementary matrix of type 1 and type 2 is symmetric. So the statement holds naturally. Let E be the elementary matrix of type 3 of adding k times of the i-th row to the j-th row. We know that E^t is also an elementary matrix of type 3. By the previous exercise we know that this kind of matrix must have determinant 1.

30. We can interchange the i-th row and the $(n+1-i)$-th row for all $i = 1, 2, \ldots \lfloor \frac{n}{2} \rfloor$[1]. Each process contribute -1 one time. So we have that

$$\det(B) = (-1)^{\lfloor \frac{n}{2} \rfloor} \det(A).$$

4.3 Properties of Determinants

1. (a) No. The elementary of type 2 has determinant other than 1.

   (b) Yes. This is Theorem 4.7.

   (c) No. A matrix is invertible if and only if its determinant is not zero.

   (d) Yes. The fact that $n \times n$ matrix A has rank n is equivalent to the fact that A is invertible and the fact that $\det(A) \neq 0$.

   (e) No. We have that $\det(A^t) = \det(A)$ by Theorem 4.8.

   (f) Yes. This is the instant result of Theorem 4.4 and Theorem 4.8.

   (g) No. It still require the condition that the determinant cannot be zero.

   (h) No. The matrix M_k is the matrix obtained from A by replacing column k of A by b.

2. Since we have the condition $\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21} \neq 0$, we can

---

[1]The symbol $\lfloor x \rfloor$ means the greatest integer m such that $m \leq x$.

use Cramer's rule and get the answer.

$$\begin{cases} x_1 &= \dfrac{\det\begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{pmatrix}}{\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}} = \dfrac{b_1 a_{22} - a_{12} b_2}{a_{11} a_{22} - a_{12} a_{21}}, \\[4ex] x_2 &= \dfrac{\det\begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix}}{\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}} = \dfrac{a_{11} b_1 - b_2 a_{21}}{a_{11} a_{22} - a_{12} a_{21}}. \end{cases}$$

3. The answer is $(x_1, x_2, x_3) = (4, -3, 0)$.

4. The answer is $(x_1, x_2, x_3) = (-1, -\frac{6}{5}, -\frac{7}{5})$.

5. The answer is $(x_1, x_2, x_3) = (-20, -48, -8)$.

6. The answer is $(x_1, x_2, x_3) = (-43, -109, -17)$.

7. The answer is $(x_1, x_2, x_3) = (42, 110, 18)$.

8. By Theorem 4.8 we know that $\det(A^t) = \det(A)$. So we can write Theorem 4.3 into column version by "The determinant of an $n \times n$ matrix is a lineaer function of each column when the remaining columns are held fixed.".

9. By Exercise 4.2.23, the determinant of an upper triangular matrix is the product of its diagonal entries. So it's invertible if and only if all its diagonal entries are nonzero.

10. If M is nilpotent, we say $M^k = O$ for some positive integer k. So we have

$$0 = \det(O) = \det(M^k) = (\det(M))^k.$$

This means that $\det(M)$ must be zero.

11. By Exercise 4.2.25 we have $\det(-M) = (-1)^n \det(M)$. By Theorem 4.8 we have $\det(M^t) = \det(M)$. So the conclusion is that

$$(-1)^n \det(M) = \det(M).$$

If n is odd, we can conclude that $\det(M) = 0$ and hence M is not invertible[2]. If n is even, we cannot say anything. For example, the matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is invertible while the matrix 2×2 zero matrix O_2 is not invertible.

12. This is an instant result by

$$1 = \det(I) = \det(QQ^t) = \det(Q)\det(Q^t) = \det(Q)^2.$$

---

[2]It should be remarked that "$-a = a$ implies $a = 0$" holds only when the field has charateristic other thant 2. In this question the field $\mathbb{C}$ has characteristic zero

13. (a) For the case $n = 1$ we have

$$\det(\bar{M}) = \bar{M}_{11} = \overline{\det}(M)$$

   . By induction, suppose that $\det(\bar{M}) = \overline{\det}(M)$ for M is a $k \times k$ matrix. For a $(n+1) \times (n+1)$ matrix M, we have

$$\det(\bar{M}) = \sum j = 1^n (-1)^{i+j} \bar{M}_{1j} \det(\tilde{\bar{M}}_{i}j)$$

$$= \sum j = 1^n (-1)^{i+j} \bar{M}_{1j} \det(\bar{\tilde{M}}_{i}j) = \overline{\det}(M).$$

   (b) This is an instant result by

$$1 = |\det(I)| = |\det(QQ^*)| = |\det(Q)||\det(Q^*)| = |\det(Q)|^2.$$

14. The set β is a basis if and only if β is an independent set of n elements. So this is equivalent to the set of columns of B is independent and hence B has rank n. And all of them are equivalent to that B is invertible and hence $\det(B) \neq 0$.

15. Two matrix A and B are similar means $A = C^{-1}BC$. So we get the conclusion

$$\det(A) = \det(C^{-1}BC) = \det(C^{-1})\det(B)\det(C) = \det(B).$$

16. Since the fact
$$1 = \det(I) = \det(A)\det(B)$$

   implies $\det(A)$ and $\det(B)$ cannot be zero, we know A is invertible.

17. By Exercise 4.2.25 we have

$$\det(AB) = (-1)^n \det(A)\det(B).$$

   This means $\det(A)$ and $\det(B)$ cannot be invertible simultaneously. So A or B is not invertible.

18. For the first case, let A be a matrix of type 2 meaning multiplying the i-th row by a scalar c. We have $\det(A) = c$ by Exercise 4.2.28. And since determinant is linear function of the i-th row when other rows are held fixed, we have

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

   For the second case, let A be a matrix of type 3 meaning adding c times the i-row to the j-th row. We have $\det(A) = 1$ by Exercise 4.2.28. And since determinant will not change when we adding some times the i-row to the j-th row, we have

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

19. Since the transpose of a lower triangular matrix is an upper triangular matrix, we have that the determinant of a lower triangular matrix is the product of all its diagonal entries.

20. We can expand the matrix by the n-th row and then by $(n-1)$-th row inductively. So we have that $\det(M) = \det(A)$. Similarly, if we expand the matrix below by the first row and the second row inductively, we get the identity

$$\det\begin{pmatrix} I & B \\ O & D \end{pmatrix} = \det(D).$$

21. First, if C is not invertible, the set of row vectors of C is not independent. This means the set of row vectors $\begin{pmatrix} O & C \end{pmatrix}$ is also not independent. So it's impossible that M has n independent rows and hence it's impossible that M is invertible. The conclusion is that if C is not invertible, we have

$$\det(A)\det(C) = \det(A)0 = 0 = \det(M).$$

Second, if C is invertible, we have the identity

$$\begin{pmatrix} I & O \\ O & C^{-1} \end{pmatrix}\begin{pmatrix} A & B \\ O & C \end{pmatrix} = \begin{pmatrix} A & B \\ O & I \end{pmatrix}.$$

So we get the identity

$$\det(C^{-1})\det(M) = \det(A)$$

and hence

$$\det(M) = \det(A)\det(C).$$

22. (a) We have that

$$\begin{cases} T(1) & = & 1e_1 + 1e_2 + \cdots + 1e_{n+1} \\ T(x) & = & c_0 e_1 + c_1 e_2 + \cdots + c_n e_{n+1} \\ \vdots & = & \vdots \\ T(x^n) & = & c_0^n e_1 + c_1^n e_2 + \cdots + c_n^n e_{n+1}, \end{cases}$$

where $\{e_1, e_2, \ldots, e_{n+1}\}$ is the standard basis for $\mathbb{F}^{n+1}$. So we get the desired conclusion.

(b) By Exercise 2.4.22 T is isomorphism and hence invertible. So the matrix M is also invertible and hence $\det(M) \neq 0$.

(c) We induction on n. For $n = 1$, we have $\det\begin{pmatrix} 1 & c_0 \\ 1 & c_1 \end{pmatrix} = (c_1 - c_0)$. Suppose the statement of this question holds for $n = k - 1$, consider the case for $n = k$. To continue the proof, we remark a fomula first below.

$$x^k - y^k = (x-y)(x^{k-1} + x^{k-2}y + \cdots + y^{k-1})$$

For brevity we write

$$p(x, y, k) = x^{k-1} + x^{k-2}y + \cdots + y^{k-1}.$$

Now to use the induction hypothesis we can add -1 time the first row to all other rows without changing the determinant.

$$\det \begin{pmatrix} 1 & c_0 & c_0^2 & \cdots & c_0^n \\ 1 & c_1 & c_1^2 & \cdots & c_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & c_n & c_n^2 & \cdots & c_n^n \end{pmatrix}$$

$$= \det \begin{pmatrix} 1 & c_0 & c_0^2 & \cdots & c_0^n \\ 0 & c_1 - c_0 & (c_1 - c_0)p(c_1, c_0, 2) & \cdots & (c_1 - c_0)p(c_1, c_0, n) \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & c_n - c_0 & (c_n - c_0)p(c_n, c_0, 2) & \cdots & (c_n - c_0)p(c_n, c_0, n) \end{pmatrix}$$

$$= \prod_{j=1}^{n} (c_j - c_0) \det \begin{pmatrix} 1 & p(c_1, c_0, 2) & \cdots & p(c_1, c_0, n) \\ \vdots & \vdots & & \vdots \\ 1 & p(c_n, c_0, 2) & \cdots & p(c_n, c_0, n) \end{pmatrix}$$

Now we write $e_i = (c_1^i, c_2^i, \ldots, c_n^i)^t$ for $i = 0, 1, \ldots n - 1$. So the determinant of the last matrix in the equality above can be written as

$$\det \begin{pmatrix} e_0 & e_1 + c_0 e_0 & e_2 + c_0 e_1 + c_0^2 e_1 & \cdots & e_{n-1} + c_0 e_{n-2} + \cdots + c_0^{n-1} e_0 \end{pmatrix}$$

$$= \det \begin{pmatrix} e_0 & e_1 & e_2 + c_0 e_1 & \cdots & e_{n-1} + c_0 e_{n-2} + \cdots \end{pmatrix}$$

$$= \det \begin{pmatrix} e_0 & e_1 & e_2 & \cdots & e_{n-1} + c_0 e_{n-2} + \cdots \end{pmatrix}$$

$$= \cdots = \det \begin{pmatrix} e_0 & e_1 & e_2 & \cdots & e_{n-1} \end{pmatrix}.$$

And by induction hypothesis, the value of it would be

$$\prod_{1 \le i \le j \le n} (c_j - c_i).$$

Combine two equalities above we get the desired conclusion.

23. (a) We prove that $\mathrm{rank}(A) \ge k$ first. If $\mathrm{rank}(A) = n$, then the matrix A has nonzero determinant and hence the interger k should be n. Now if $\mathrm{rank}(A) = r < n$, we prove by contradiction. If $k > \mathrm{rank}(A)$, we can find a $k \times k$ submatrix B such that the determinant of it is not zero. This means the set

$$S = \{v_1, v_2, \ldots v_k\}$$

of columns of B is independent. Now consider the $k \times n$ submatrix C of A obtained by deleting those rows who were deleted when we construct B. So S is a subset of the set of columns of C. This means

$$k \ge \mathrm{rank}(C) \le \min\{n, k\} = k$$

96

and hence $\operatorname{rank}(C) = k$. So this also means that the set of the k rows of C independent. And thus the matrix A contains k independent rows and hence $\operatorname{rank}(A) \geq k$, a contradiction.

Conversely, we construct a $r \times r$ submatrix of A, where r is $\operatorname{rank}(A)$, to deduce that $\operatorname{rank}(A) \leq k$. Since rank of A is r, we have r independent rows, say $u_1, u_2, \ldots, u_r$. Let D be the $r \times n$ submatrix such that the i-th row of D is u_i. Since the set of rows of D is independent, we have that

$$r \leq \mathrm{D} \leq \min\{r, n\} = r$$

and hence $\operatorname{rank}(D) = r$. Similarly we have $w_1, w_2, \ldots, w_r$ to be the r independent columns of D. And si-similarly we can construct a $r \times r$ matrix E such that the i-th column of E is w_i. Since E is a $r \times r$ matrix with r independent rows, we have $\operatorname{rank}(E) = r$. This complete the proof.

(b) See the second part of the previous exercise.

24. We use induction to claim that

$$\det(A + tI) = t^n + \sum_{i=n-1}^{0} a_i t^i.$$

For $n = 1$, it's easy to see that $\det(A + tI) = t + a_0$. Suppose that the statement holds for $n = k - 1$, consider the case for $n = k$. We can expand the matrix and get

$$\det \begin{pmatrix} t & 0 & 0 & \cdots & 0 & a_0 \\ -1 & t & 0 & \cdots & 0 & a_1 \\ 0 & -1 & t & \cdots & 0 & a_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & a_{k-1}+t \end{pmatrix}$$

$$= t \det \begin{pmatrix} t & 0 & \cdots & 0 & a_1 \\ -1 & t & \cdots & 0 & a_2 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & -1 & a_{k-1} \end{pmatrix} + (-1)^{1+k} a_0 \det \begin{pmatrix} -1 & t & \cdots & 0 \\ 0 & t & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t \end{pmatrix}$$

$$= t\left(t^{k-1} + \sum_{i=k-2}^{0} a_{i+1} t^i\right) + (-1)^{1+k} a_0 (-1)^{k-1}$$

$$= t^n + \sum_{i=n-1}^{0} a_i t^i.$$

25. (a) Just expand along the k-th column.

(b) It's better not to use a great theorem, such as Cramer's rule, to kill a small problem. We check each entry one by one. First, we have that

$$\sum_{k=1}^{n} A_{jk} c_{jk} = \det(A)$$

97

and so the j-th entry of the left term is $\det(A)$. Second, for $i \neq j$, we construct a matrix B_i by replacing the j-th row of A by the i-th row of A. Since B_i has two identity rows, we have that $\det(B_i) = 0$ for all $i \neq j$. Now we can calculate that

$$\sum_{k=1}^{n} A_{ik} c_{jk} = \sum_{k=1}^{n} B_{jk} c_{jk} = \det(B) = 0,$$

for all $i \neq j$. So we get the desired conclusion.

(c) Actually this matrix C is the classical adjoint of matrix A defined after this exercise. And this question is an instant result since

$$AC = A \begin{pmatrix} c_{11} & c_{21} & \cdots & c_{n1} \\ c_{12} & c_{22} & \cdots & c_{n2} \\ \vdots & \vdots & & \vdots \\ c_{1n} & c_{2n} & \cdots & c_{nn} \end{pmatrix}$$

$$= A \begin{pmatrix} \det(A) & 0 & \cdots & 0 \\ 0 & \det(A) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \det(A) \end{pmatrix}$$

by the previous exercise.

(d) If $\det(A) \neq 0$, then we know A is invertible. So we have

$$A^{-1} = A^{-1} A [\det(A)]^{-1} C = [\det(A)]^{-1} C.$$

26. (a) We have that
$$c_{11} = (-1)^2 \det((\tilde{A})_{11}) = A_{22},$$
$$c_{12} = (-1)^3 \det((\tilde{A})_{12}) = -A_{21},$$
$$c_{21} = (-1)^3 \det((\tilde{A})_{21}) = -A_{12},$$
$$c_{22} = (-1)^4 \det((\tilde{A})_{22}) = A_{11}.$$

So the adjoint of matrix A is

$$\begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}.$$

(b) The adjoint of that matrix is

$$\begin{pmatrix} 16 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 16 \end{pmatrix}.$$

(c) The adjoint of that matrix is

$$\begin{pmatrix} 10 & 0 & 0 \\ 0 & -20 & 0 \\ 0 & 0 & -8 \end{pmatrix}.$$

(d) The adjoint of that matrix is

$$\begin{pmatrix} 20 & -30 & 20 \\ 0 & 15 & -24 \\ 0 & 0 & 12 \end{pmatrix}.$$

(e) The adjoint of that matrix is

$$\begin{pmatrix} -3i & 0 & 0 \\ 4 & -1+i & 0 \\ 10+16i & -5-3i & 3+3i \end{pmatrix}.$$

(f) The adjoint of that matrix is

$$\begin{pmatrix} 6 & 22 & 12 \\ 12 & -2 & 24 \\ 21 & -38 & -27 \end{pmatrix}.$$

(g) The adjoint of that matrix is

$$\begin{pmatrix} 18 & 28 & -6 \\ -20 & -21 & 37 \\ 48 & 14 & -16 \end{pmatrix}.$$

(h) The adjoint of that matrix is

$$\begin{pmatrix} -i & -8+i & -1+2i \\ 1-5i & 9-6i & -3i \\ -1+i & -3 & 3-i \end{pmatrix}.$$

27. (a) If A is not invertible, we have $AC = [\det(C)]I = O$. It's impossible that C is invertible otherwise $A = C^{-1}O = O$. But the adjoint of the zero matrix O is also the zero matrix O, which is not invertible. So we know that in this case C is not invertible and hence $\det(C) = 0 = [\det(A)]^{n-1}$. Next, if A is invertible, we have, by Exercise 4.3.25(c) that

$$\det(A)\det(C) = \det([\det(A)]I) = [\det(A)]^n$$

So we know that

$$\det(C) = [\det(A)]^{n-1}$$

since $\det(A) \neq 0$.

(b) This is because

$$\det(\tilde{A}^t{}_{ij}) = \det(\tilde{A}^t_{ji}) = \det(\tilde{A}).$$

(c) If A is an invertible upper triangular matrix, we claim that $c_{ij} = 0$ for all i, j with $i > j$. For every i, j with $i > j$, we know that

$$c_{ij} = (-1)^{i+j} \det(\tilde{A}_{ij}).$$

But $\tilde{A}_{ij}$ is an upper triangular matrix with at least one zero diagonal entry if $i > j$. Since determinant of an upper triangular matrix is the product of all its diagonal entries. We know that for $i > j$ we have $\det(\tilde{A}_{ij}) = 0$ and hence $c_{ij} = 0$. With this we know that the adjoint of A is also a upper triangular matrix.

28. (a) For brevity, we write

$$v(y(t)) = (y(t), y'(t), \cdots, y^{(n)}(t))^t$$

and

$$v_i(t) = (y(t), y'(t), \cdots, y^{(n)}(t))^t.$$

Since the defferential operator is linear, we have

$$v((x + cy)(t)) = v(x(t)) + cv(y(t)).$$

Now we have that

$$[T(x + cy)](t) = \det\big(v((x + cy)(t)) \quad v_1(t) \quad v_2(t) \quad \cdots \quad v_n(t)\big)$$

$$= \det\big(v(x(t)) + cv(y(t)) \quad v_1(t) \quad v_2(t) \quad \cdots \quad v_n(t)\big)$$

$$= [T(x)](t) + c[T(y)](t)$$

since determinant is a linear function of the first column when all other columns are held fixed.

(b) Since $N(T)$ is a space, it enough to say that $y_i \in N(T)$ for all i. But this is easy since

$$[T(y)](t) = \det(v_i(t), v_1(t), \ldots, v_n(t)) = 0.$$

The determinant is zero since the matrix has two identity columns.

4.4   Summary—Important Facts about Determinants

1. With the help of Theorem 4.8 the statement in row version or column version is equivalent. So we won't mention it again and again.

(a) Yes. It's Exercise 4.3.8.

(b) Ur...No! It's wise to check whether there are any two identity rows or two identity columns first. How do we know what is wise?

(c) Yes. See the Corollary after Theorem 4.4.

(d) No. The determinant should be multiplied by −1.

(e) No. The scalar cannot be zero.

(f) Yes. Watch Theorem 4.6.

(g) Yes. This is Exercise 4.2.23.

(h) No. Read Theorem 4.8.

(i) Yes. Peruse Theorem 4.7.

(j) Yes. Glance the Corollary after Theorem 4.7.

(k) Yes. Gaze the Corollary after Theorem 4.7.

2. (a) The determinant is 22.

(b) The determinant is −29.

(c) The determinant is $2 - 4i$.

(d) The determinant is $30i$.

3. (a) The determinant should be −12.

(b) The determinant should be −13.

(c) The determinant should be −12.

(d) The determinant should be −13.

(e) The determinant should be 22.

(f) The determinant should be $4 + 2i$.

(g) The determinant should be −3.

(h) The determinant should be 154.

4. (a) The determinant should be 0.

(b) The determinant should be 36.

(c) The determinant should be −49.

(d) The determinant should be 10.

(e) The determinant should be $-28 - i$.

(f) The determinant should be $17 - 3i$.

(g) The determinant should be 95.

(h) The determinant should be −100.

5. See Exercise 4.3.20. Remember to write it down clearly one more time. The given proof there is simplistic.

6. Ur...look Exercise 4.3.21.

4.5 A Characterization of the Determinant

1. (a) No. For example, we have $\det(2I) \neq 2\det(I)$ when I is the 2×2 identity matrix.

   (b) Yes. This is Theorem 4.3.

   (c) Yes. This is Theorem 4.10.

   (d) No. Usually it should be $\delta(B) = -\delta(A)$.

   (e) No. Both the determinant and the zero function, $\delta(A) = 0$, are n-linear function.

   (f) Yes. Let $v_1, u_1, u_2, \ldots, u_n$ are row vectors and c are scalar. We have

$$\delta\begin{pmatrix} u_1 + cv_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = 0 = 0 + c \cdot 0 = \delta\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + c\delta\begin{pmatrix} v_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}.$$

   The cases for other rows are similar.

2. A 1-linear function is actually a linear function. We can deduce that

$$\delta(x) = x\delta(1) = ax,$$

   where a is defined by (1). So all the functions must be in the form $\delta(x) = ax$.

3. It's not a 3-linear function. We have that

$$\delta\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = k \neq 2\delta(I_3) = 2k.$$

4. It's not a 3-linear function. We have that

$$\delta\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 1 \neq 2\delta(I_3) = 2.$$

5. It's a 3-linear function. We have that when the second and the third rows are held fixed, the function would be

$$\delta\begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} = (A_{11}, A_{12}, A_{13}) \cdot (A_{23}A_{32}, 0, 0),$$

   a inner product function. So δ is linear for the first row. Similarly we can write

$$\delta\begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}$$

$$= (A_{21}, A_{22}, A_{23}) \cdot (0, 0, A_{11}A_{32}) = (A_{31}, A_{32}, A_{33}) \cdot (0, A_{11}A_{23}, 0).$$

So δ is a 3 linear function.

6. It's not a 3-linear function. We have that

$$\delta \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 4 \neq= 2\delta(I_3) = 6.$$

7. It's a 3 linear function. We could write

$$\delta \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} = (A_{11}, A_{12}, A_{13}) \cdot (A_{21}A_{32}, 0, 0)$$

$$= (A_{21}, A_{22}, A_{23}) \cdot (A_{11}A_{32}, 0, 0) = (A_{31}, A_{32}, A_{33}) \cdot (0, A_{11}A_{21}, 0)$$

and get the result.

8. It's not a 3-linear function. We have that

$$\delta \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 1 & 1 & 0 \end{pmatrix} = 1 \neq= 2\delta \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} = 2.$$

9. It's not a 3-linear function. We have that

$$\delta \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 4 \neq= 2\delta(I_3) = 2.$$

10. It's a 3 linear function. We could write

$$\delta \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} = (A_{11}, A_{12}, A_{13}) \cdot (A_{22}A_{33} - A + 21A_{32}, 0, 0)$$

$$= (A_{21}, A_{22}, A_{23}) \cdot (-A_{11}A_{32}, A_{11}A_{33}, 0) = (A_{31}, A_{32}, A_{33}) \cdot (0, A_{11}A_{21}, A_{11}A_{22})$$

and get the result.

11. **Corollary 2.** Since δ is n-linear, we must have $\delta(A) = 0$ if A contains one zero row. Now if M has rank less than n, we know that the n row vectors of M are dependent, say $u_1, u_2, \ldots, u_n$. So we can find some vector u_i who is a linear combination of other vectors. We write

$$u_i = \sum j \neq i a_j u_j.$$

By Corollary 1 after Theorem 4.10 we can add $-a_j$ times the j-th row to the i-th row without changing the value of δ. Let M' be the matrix obtained from M by doing this processes. We know M' has one zero row, the i-th row, and hence $\delta(M_\delta(M') = 0$.

Corollary 3 We can obtain E_1 from I by interchanging two rows. By Theorem 4.10(a) we know that $\delta(E_1) = -\delta(I)$. Similarly we can obtain E_2 from I by multiplying one row by a scalar k. Since δ is n-linear we know that $\delta(E_2) = k\delta(I)$. Finally, we can obtain E_3 by adding k times the i-th row to the j-th row. By Corollary 1 after Theorem 4.10 we know that $\delta(E_3) = \delta(I)$.

12. If A is not full-rank, we have that AB will be not full-rank. By Corollary 3 after Theorem 4.10 we have

$$\delta(AB) = 0 = \delta(A)\delta(B).$$

If A is full-rank and so invertible, we can write $A = E_s \cdots E_2 E_1$ as product of elementary matrices. Assuming the fact, which we will prove later, that

$$\delta(EM) = \delta(E)\delta(M)$$

for all elementary matrix E and all matrix M holds, we would have done since
$$\delta(AB) = \delta(E_s \cdots E_2 E_1 B) = \delta(E_s)\delta(E_{k-1} \cdots E_2 E_1 B)$$

$$= \cdots = \delta(E_s) \cdots \delta(E_2)\delta(E_1)\delta(B)$$

$$= \delta(E_s \cdots E_2 E_1)\delta(B) = \delta(A)\delta(B).$$

So now we prove the fact. First, if E is the elementary matrix of type 1 meaning interchangine the i-th and the j-th rows, we have EM is the matrix obtained from M by interchanging the i-th and the j-th rows. By Theorem 4.10(a) we know that

$$\delta(EM) = -\delta(M) = -\delta(I)\delta(M) = \delta(E)\delta(M).$$

Second, if E is the elementary matrix of type 2 meaning multiplying the i-th row by a scalar k, we have EM is the matrix obtained from M by multiplying the i-th row by scalar k. Since the function δ is n-linear, we have
$$\delta(EM) = k\delta(M) = k\delta(I)\delta(M) = \delta(E)\delta(M).$$

Finally, if E is the elementary matrix of type 3 meaning adding k times the i-th row to the j-th row, we have EM is the matrix obtained from M by adding k times the i-th row to the j-th row. By Corollary 1 after Theorem 4.10, we have

$$\delta(EM) = \delta(M) = \delta(I)\delta(M) = \delta(E)\delta(M).$$

This complete the proof.

13. Since the fact $\det(A^t) = \det(A)$ and det is a 2-linear function, the result is natural.

14. We could write

$$\delta \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = (A_{11}, A_{12}) \cdot (A_{22}a + A_{21}b, A_{22}c + A_{21}d)$$

$$= (A_{21}, A_{22}) \cdot (A_{11}b + A_{12}d, A_{11}a + A_{12}c)$$

and get the desired result since inner product function is linear.

For the converse, fixed one n-linear function δ and let

$$a = \delta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, b = \delta \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix},$$

$$c = \delta \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, d = \delta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Now we must have

$$\delta \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = A_{11}\delta \begin{pmatrix} 1 & 0 \\ A_{21} & A_{22} \end{pmatrix} + A_{12}\delta \begin{pmatrix} 0 & 1 \\ A_{21} & A_{22} \end{pmatrix}$$

$$= A_{11}(A_{21}\delta \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + A_{22}\delta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}) + A_{12}(A_{21}\delta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + A_{22}\delta \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix})$$

$$= A_{11}A_{22}a + A_{11}A_{21}b + A_{12}A_{22}c + A_{12}A_{21}d.$$

15. Wait

16. Fixed an alternating n-linear function δ. Let k be the value of $\delta(I)$. We want to chaim that
$$\delta(M) = k \det(M).$$

First we know that if M has rank less than n, then $\delta(M) = 0 = \det(M)$ by Corollary 2 after Theorem 4.10. So the identity holds. Second if M is full-rank, we can write $M = E_s \cdots E_2 E_1 I$ as product of elementary matrices and identity matrix I. And it's lucky now I can copy and paste the text in Exercise 4.5.12.

This time we will claim that

$$\delta(EA) = \det(E)\delta(A)$$

for all elementary matrix E and all matrix A. First, if E is the elementary matrix of type 1 meaning interchangine the i-th and the j-th rows, we have EM is the matrix obtained from A by interchanging the i-th and the j-th rows. By Theorem 4.10(a) we know that

$$\delta(EM) = -\delta(A) = \det(E)\delta(A).$$

Second, if E is the elementary matrix of type 2 meaning multiplying the i-th row by a scalar k, we have EM is the matrix obtained from M by

multiplying the i-th row by scalar k. Since the function δ is n-linear, we have

$$\delta(EM) = k\delta(A) = \det(E)\delta(A).$$

Finally, if E is the elementary matrix of type 3 meaning adding k times the i-th row to the j-th row, we have EM is the matrix obtained from M by adding k times the i-th row to the j-th row. By Corollary 1 after Theorem 4.10, we have

$$\delta(EM) = \delta(A) = \det(E)\delta(A).$$

This complete the proof since

$$\delta(M) = \delta(E_s\cdots E_2 E_1 I) = \det(E_s)\cdots\det(E_2)\det(E_1)\delta(I)$$

$$= k\det(E_s)\cdots\det(E_2)\det(E_1) = k\det(M).$$

17. Recall the definition

$$(\delta_1 + \delta_2)(A) = \delta_1(A) + \delta_2(A)$$

and

$$(k\delta)(A) = k\delta(A)$$

. For brevity, we write δ' for $\delta_1 + \delta_2$ and δ'' for $k\delta$. Now prove that both δ' and δ'' is n-linear. Check that

$$\delta'\begin{pmatrix} u_1 + cv_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \delta_1\begin{pmatrix} u_1 + cv_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + \delta_2\begin{pmatrix} u_1 + cv_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$

$$= \delta_1\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + c\delta_1\begin{pmatrix} v_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + \delta_2\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + \delta_2\begin{pmatrix} v_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$

$$= \delta'\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + c\delta'\begin{pmatrix} v_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}.$$

Also check that

$$\delta''\begin{pmatrix} u_1 + cv_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = k\delta\begin{pmatrix} u_1 + cv_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$

$$= k\delta\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + ck\delta\begin{pmatrix} v_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \delta''\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + c\delta''\begin{pmatrix} v_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}.$$

So both δ' and δ'' is linear function for the first row when other rows are held fixed. For the cases on other rows are similar.

18. Let the zero element be the zero n-linear function, $\delta(A) = 0$. Thus it can be checked that all the properties of vector space hold by the properties of field.

19. If M has two identical rows, then the matrix M' obtained from M by interchanging the two rows would be the same. So we have

$$\delta(M) = \delta(M') = -\delta(M).$$

When $\mathbb{F}$ does not have characteristic two, the equality above means $\delta(M) = 0$.

20. Let δ be the 2-linear function in Exercise 4.5.15 with $a = b = c = d = 1$. Thus we have that

$$\delta\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ac + ad + bc + bd = \delta\begin{pmatrix} c & d \\ a & b \end{pmatrix} = -\delta\begin{pmatrix} c & d \\ a & b \end{pmatrix}.$$

The final equality holds since $\mathbb{F}$ has characteristic two. But now we have

$$\delta\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = 1 \neq 0.$$

Chapter 5

Diagonalization

5.1 Eigenvalues and Eigenvectors

1. (a) No. For example, the identity mapping I_2 has two eigenvalues 1, 1.

   (b) Yes. If we have $(A - \lambda I)v = 0$, we also have

   $$(A - \lambda I)(cv) = c(A - \lambda I)v = 0$$

   for all $c \in \mathbb{R}$. Note that this skill will make the statement false when the field is finite.

   (c) Yes. For example, the matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ means the rotation through the angle $\frac{\pi}{2}$. And the matrix has no real eigenvalue and hence no real eigenvector.

   (d) No. See definition.

   (e) No. For the matrix I_2, the vectors $(1,0)$ and $(2,0)$ are all eigenvectors but they are parallel.

   (f) No. The matrix $\begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}$ has only two eigenvalues 1 and -1. But the sum 0 is not an eigenvalue of the same matrix.

   (g) No. Let P be the space of all polynomial and T be the identity mapping from P to P. Thus we know 1 is an eigenvalue of T.

   (h) Yes. That a matrix A is similar to a diagonal matrix D means there is some invertible matrix P such that $P^{-1}AP = D$. Since P is invertible, $P = [I]_\alpha^\beta$ for some basis α ,where β is the standard basis for $\mathbb{F}^n$. So the first statement is equivalent to that A is diagonalizable. And the desired result comes from Theorem 5.1.

   (i) Yes. If A and B are similar, there is some invertible matrix P such that $P^{-1}AP = B$. If $Av = \lambda v$, we have

   $$B(P^{-1}v) = \lambda P^{-1}v.$$

And if $Bv = \lambda v$, we have $A(Pv) = Pv$.

(j) No. It usually false. For example, the matrices $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ and $\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$ are similar since

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

But the eigenvector $(1,0)$ of the first matrix is not a eigenvector of the second matrix.

(k) No. The vectors $(1,0)$ and $(0,1)$ are eigenvectors of the matrix $\begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}$. But the sum of them $(1,1)$ is not an eigenvector of the same matrix.

2. Compute $[T]_\beta$ as what we did in previous chapters. If that matrix is diagonal, then β is a basis consisting of eigenvectors.

(a) No. $[T]_\beta = \begin{pmatrix} 0 & 2 \\ -1 & 0 \end{pmatrix}$.

(b) Yes. $[T]_\beta = \begin{pmatrix} -2 & 0 \\ 0 & -3 \end{pmatrix}$.

(c) Yes. $[T]_\beta = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

(d) No. $\begin{pmatrix} 0 & 0 & 3 \\ 0 & -2 & 0 \\ -4 & 0 & 0 \end{pmatrix}$.

(e) No. $[T]_\beta = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$.

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

3. Calculate the characteristic polynomial of A and find the zeroes of it to solve (i). Find the nonzero vector v such that $(A - \lambda I)v = 0$ to solve (ii). For (iii) and (iv) just follow the direction given by the textbook.

(a) The characteristic polynomial is $t^2 - 3t - 4$ and the zeroes of it are 4 and -1. For eigenvalue 4, we may solve (A-4I)x=0. There are infinite

solutions. Just pick one from them, say $(2,3)$. Similarly we can find the eigenvector corresponding to -1 is $(1,-1)$. Pick

$$\beta = \{(2,3),(1,-1)\}$$

and

$$Q = [I]_\beta^\alpha = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix},$$

where α is the standard basis for $\mathbb{R}^2$. Then we know that

$$D = Q^{-1}AQ = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}.$$

(b) The characteristic polynomial is $-(t-3)(t-2)(t-1)$ with 3 zeroes 1, 2, and 3. The corresponding eigenvectors are $(1,1,-1)$, $(1,-1,0)$, and $(1,0,-1)$. The set of these three vectors are the desired basis. And we also have

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

and

$$D = Q^{-1}AQ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

(c) The characteristic polynomial is $(t-1)(t+1)$ with two zeroes -1 and 1. The corresponding eigenvectors are $(1,-i-1)$ and $(1,-i+1)$. The set of these two vectors are the desired basis. And we also have

$$Q = \begin{pmatrix} 1 & 1 \\ -i-1 & -i+1 \end{pmatrix}$$

and

$$D = Q^{-1}AQ = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

(d) The characteristic polynomial is $-(t-1)^2 t$ with zeroes 0, 1, and 1. The corresponding eigenvectors are $(1,4,2)$, $(1,0,1)$, and $(0,1,0)$. The set of these three vectors are the desired basis. And we also have

$$Q = \begin{pmatrix} 1 & 1 & 0 \\ 4 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}$$

and

$$D = Q^{-1}AQ = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

4. Follow the process of the previous exercise.

(a) $\beta = \{(3,5),(1,2)\}$ and $[T]_\beta = \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}$.

(b) $\beta = \{(2,0,-1),(1,2,0),(1,-1,-1)\}$ and $[T]_\beta = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

(c) $\beta = \{(1,-2,-2),(2,0,-1),(0,2,1)\}$ and $[T]_\beta = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

(d) $\beta = \{2x+3, x+2\}$ and $[T]_\beta = \begin{pmatrix} -3 & 0 \\ 0 & -2 \end{pmatrix}$.

(e) $\beta = \{x-3, 4x^2-13x-3, x+1\}$ and $[T]_\beta = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$.

(f) $\beta = \{x^3-8, x^2-4, x-2, x\}$ and $[T]_\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$.

(g) $\beta = \{(1, x-1, 3x^2-2, 2x^3+6x-7\}$ and $[T]_\beta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$.

(h) $\beta = \{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\}$ and $[T]_\beta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

(i) $\beta = \{\begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}\}$ and $[T]_\beta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

(j) $\beta = \{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\}$ and $[T]_\beta = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

5. By definition, that v is an eigenvector of T corresponding to λ means $Tv = \lambda v$ and $v \neq 0$. Hence we get

$$Tv - \lambda v = (T - \lambda I)v = 0$$

and $v \in N(T - \lambda I)$. Conversely, if $v \neq 0$ and $v \in N(T - \lambda I)$, we have $Tv = \lambda v$. It's the definition of eigenvector.

6. We know that
$$Tv = \lambda v$$
is equivalent to
$$[T]_\beta [v]_\beta = \lambda [v]_\beta.$$

7. (a) We have that
$$[T]_\beta = [I]_\gamma^\beta [T]_\gamma [I]_\beta^\gamma$$
and
$$([I]_\gamma^\beta)^{-1} = [I]_\beta^\gamma.$$
So we know that
$$\det([T]_\beta) = \det([I]_\gamma^\beta) \det([T]_\gamma) \det([I]_\beta^\gamma) = \det([T]_\gamma).$$

(b) It's the instant result from Theorem 2.18 and the Corollary after Theorem 4.7.

(c) Pick any ordered basis β and we have
$$\det(T^{-1}) = \det([T^{-1}]_\beta) = \det(([T]_\beta)^{-1}) = \det([T]_\beta)^{-1} = \det(T)^{-1}.$$
The second and the third equality come from Theorem 2.18 and the Corollary after Theorem 4.7.

(d) By definition,
$$\det(TU) = \det([TU]_\beta) = \det([T]_\beta [U]_\beta)$$
$$= \det([T]_\beta) \det([U]_\beta) = \det(T) \det(U).$$

(e) We have
$$\det(T - \lambda I_V) = \det([T - \lambda I]_\beta) = \det([T]_\beta - \lambda [I]_\beta) = \det([T]_\beta - \lambda I).$$

8. (a) By previous exercises, we have that T is invertible if and only if $\det(T) \neq 0$. And the fact $\det(T) \neq 0$ is equivalent to the fact $N(T - 0I) = \{0\}$, which is equivalent to that zero is not an eigenvalue of T by Theorem 5.4.

(b) Since $T = (T^{-1})^{-1}$, it's enough to prove only one side of the statement. If λ an eigenvalue with eigenvector v, we have $Tv = \lambda v$ and so $T^{-1}v = \lambda^{-1}v$. This means λ^{-1} is an eigenvalue of T^{-1}.

(c) **(a)** A matrix M is invertible if and only if 0 is not an eigenvalue of M.

(b) Let M be an invertible matrix. We have λ is an eigenvalue of M if and only if λ^{-1} is an eigenvalue of M^{-1}.

First, if M is invertible, then there's no vector v such that $Mv = 0v = 0$. So 0 is not an eigenvalue of M. If 0 is not an eigenvalue of M, then $v = 0$ is the only vector sucht that $Mv = 0$. This means that M is injective and so invertible since M is square. Second, it's enough to prove one side of that statement since $M = (M^{-1})^{-1}$. And if we have $Mv = \lambda v$, then we have $M^{-1}v = \lambda^{-1}v$.

9. This is directly because the determinant of an upper triangular matrix is the product of it's diagonal entries.

10. (a) What did the author want to ask? Just calculate it!

    (b) Just calculate it and get the answer $(\lambda - t)^n$, where n is the dimension of V.

    (c) We already have that for any ordered basis β, $[\lambda I_V]_\beta = \lambda I$, a diagonal matrix. So it's diagonalizable. By the previous exercise we also have that the only eigenvalue is the zero of the characteristic polynomial, λ.

11. (a) It's just because
    $$A = B^{-1}\lambda I B = \lambda I$$
    for some invertible matrix B.

    (b) Let M be the matrix mentioned in this question and λ be that only eigenvalue of M. We have that the basis to make M diagonalizable if consisting of vectors v_i such that $(M - \lambda I)v_i = 0$. This means $(M - \lambda I)v = 0$ for every v since $\{v_i\}$ forms a basis. So M must be λI.

    (c) It's easy to see that 1 is the only eigenvalue of the matrix. But since nullity of
    $$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$
    is one, we can't find a set of two vector consisting of eigenvectors such that the set is independent. By Theorem 5.1, the matrix is not diagonalizable.

12. (a) Use the fact if $A = P^{-1}BP$, then
    $$\det(A - \lambda I) = \det(P^{-1}(A - \lambda I)P)$$
    $$= \det(P^{-1}AP - \lambda P^{-1}IP) = \det(B - \lambda I).$$

    (b) Use the result of the previous exercise and the fact that each matrix representation of one linear operator is pairwisely similar to each other.

13. (a) Since the diagram is commutative, we know that
    $$T(v) = \phi_\beta^{-1}(T(\phi_\beta(v))) = \phi_\beta^{-1}(\lambda \phi_\beta(v)) = \lambda v.$$

    (b) One part of this statement has been proven in the previous exercise. If $\phi_\beta^{-1}(y)$ is an eigenvector of T corresponding to λ, we have
    $$Ay = \phi_\beta(T(\phi_\beta^{-1}(y))) = \phi_\beta(\lambda \phi_\beta(y)) = \lambda y.$$

14. Use the fact

$$\det(A - \lambda I) = \det((A - \lambda I)^t) = \det(A^t - \lambda I).$$

15. (a) If we have $T(v) = \lambda v$ for some v, then we also have that

$$T^m(v) = T^{m-1}(\lambda v) = \lambda T^{m-1}(v) = \cdots = \lambda^m v.$$

(b) You can just replace the character T by the character A.

16. (a) Just as the Hint, we have

$$\mathrm{tr}(P^{-1}AP) = \mathrm{tr}(PP^{-1}A) = \mathrm{tr}(A).$$

(b) We may define the trace of a linear operator on a finite-dimensional vector space to be the trace of its matrix representation. It's well-defined due to the previous exercise.

17. (a) If $T(A) = A^t = \lambda A$ for some λ and some nonzero matrix A, say $A_{ij} \neq 0$, we have

$$A_{ij} = \lambda A_{ji}$$

and

$$A_{ji} = \lambda A_{ij}$$

and so

$$A_{ij} = \lambda^2 A_{ij}.$$

This means that λ can only be 1 or -1. And these two values are eigenvalues due to the existence of symmetric matrices and skew-symmetric matrices.

(b) The set of nonzero symmetric matrices are the eigenvectors corresponding to eigenvalue 1, while the set of nonzero skew-symmetric matrices are the eigenvectors corresponding to eigenvalue -1.

(c) It could be $\{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\}.$

(d) Let E_{ij} be the matrix with its ij-entry 1 and all other entries 0. Then the basis could be

$$\{E_{ii}\}_{i=1,2,\dots,n} \cup \{E_{ij} + E_{ji}\}_{i>j} \cup \{E_{ij} - E_{ji}\}_{i>j}.$$

18. (a) If B is invertible, we have B^{-1} exists and $\det(B) \neq 0$. Now we know that

$$\det(A + cB) = \det(B)\det(B^{-1}A + cI),$$

a nonzero polynomial of c. It has only finite zeroes, so we can always find some c sucht that the determinant is nonzero.

(b) Since we know that
$$\det \begin{pmatrix} 1 & 1 \\ c & 1+c \end{pmatrix} = 1,$$

take $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$.

19. Say $A = P^{-1}BP$. Pick $V = \mathbb{F}^n$, $T = L_A$, and β be the standard basis. We can also pick γ such that $P = [I]^\beta_\gamma$. That is, γ is the set of the column vectors of P.

20. From the equation given in the question, it's easy that $f(0) = a_0$. And by definition, we also know that $f(t) = \det(A - tI)$ and so $f(0) = \det(A)$. So A is invertible if and only if $a_0 = \det(A) \neq 0$.

21. (a) We should use one fact that if B is a matrix with number of nonzero entries less than or equal to k, then we have $\det(A - tB)$ is a polynomial of t with degree less than or equal to k. To prove this, we induct on both k and the size of matrix n. If $k = 0$, we know that B is a zero matrix and $\det(A - tB)$ is constant. For $n = 1$, it's easy to see that degree of $\det(A - tB)$ is equal to 1, which will be less than or equal to k if $k \geq 1$. Suppose the hypothesis holds for $n = k - 1$. For the case $n = k$, we may expand the determinant along the first row. That is,

$$\det(A - tB) = \sum_{j=1}^{n} (-1)^{1+j} (A - tB)_{1j} \det(A \tilde{-} tB_{1j}).$$

If the first row of B is all zero, then $\det(A \tilde{-} tB_{1j})$ is a polynomial with degree less than or equal to k and $(A - tB)_{1j}$ contains no t for all j. If the first row of B is not all zero, then $\det(A \tilde{-} tB_{1j})$ is a polynomial with degree less than or equal to $k - 1$ and each $(A - tB)_{1j}$ contains t with degree at most 1. In both case, we get that $\det(A - tB)$ is a polynomial with degree less than or equal to k. Now we may induct on n to prove the original statement. For $n = 1$, we have $f(t) = A_{11} - t$. For $n = 2$, we have $f(t) = (A_{11} - t)(A_{22} - t) - A_{12}A_{21}$. Suppose the hypothesis is true for $n = k - 1$. For the case $n = k$, we expand the determinant along the first row. That is,

$$\det(A - tI) = (A_{11} - t)\det((\tilde{A} - tI)_{11}) + \sum_{j=2}^{k} (-1)^{i+j} \det((\tilde{A} - tI)_{1j}).$$

By the induction hypothesis, we know that

$$\det((\tilde{A} - tI)_{11}) = (A_{22} - t)(A_{33} - t)\cdots(A_{kk} - t) + p(t),$$

where $p(t)$ is a polynomial with degree less than or equal to $k - 3$, and

$$(-1)^{i+j} \det((\tilde{A} - tI)_{1j})$$

115

is a polynomial with degree less than or equal to $k-2$. So it becomes

$$(A_{11}-t)(A_{22}-t)\cdots(A_{nn}-t)+(A_{11}-t)p(t)+\sum_{j=2}^{n}(-1)^{i+j}\det((\tilde{A}-tI)_{1j},$$

in which the summation of the second term and the third term is a polynomial with degree less than or equal to $n-1$.

(b) By the previous exercise, we know that the coefficient of t^{n-1} comes from only the first term

$$(A_{11}-t)(A_{22}-t)\cdots(A_{nn}-t)$$

and it would be

$$(-1)^{n-1}\sum A_{ii}=\operatorname{tr}(A).$$

22. (a) Use the notation x and λ that in this question. In Exercise 5.1.15(a) we already have
$$T^m(x)=\lambda^m x.$$

And it's also easy to see that if $Ax=\lambda_1 x$ and $Bx=\lambda_2 x$, we'll have

$$(A+B)x=\lambda_1+\lambda_2.$$

So we get the desired result.

(b) Shit! It's not funny! Just change T into A.

(c) Just calculate that
$$g(A)=\begin{pmatrix}14 & 10\\15 & 19\end{pmatrix}$$

and

$$g(A)\begin{pmatrix}2\\3\end{pmatrix}=29\begin{pmatrix}2\\3\end{pmatrix}.$$

Also check that $\lambda=4$ and $g(4)=29$.

23. If T is diagonalizable, there is a basis β consisting of eigenvectors. By the previouse exercise, we know that if $T(x)=\lambda x$,

$$f(T)(v)=f(\lambda)v=0v=0.$$

This means that for each $v\in\beta$, we have $f(T)(v)=0$. Since β is a basis, $f(T)=T_0$.

24. (a) The coefficient of t^n comes from the first term of that equality in Exercise 5.1.21(a). So it's $(-1)^n$.

(b) A polynomial of degree n has at most n zeroes.

25. Where is the Corollaries?

116

26. By Exercise 5.1.20 and Exercise 5.1.21(b) and Theorem 5.3(b), we know the characteristic polynomial must be

$$t^2 - \mathrm{tr}(A)t + \det(A).$$

And the coefficient could be 0 or 1. So there are 4 possible characteristic polynomials. By checking that all of them are achievable, we get the answer are 4.

5.2 Diagonalizability

1. (a) No. The matrix
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

has only two distinct eigenvalues but it's diagonalizable.

(b) No. The vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ are both the eigenvectors of the matrix

$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ corresponding the same eigenvalue 1.

(c) No. The zero vector is not.

(d) Yes. If $x \in E_{\lambda_1} \cap E_{\lambda_2}$, we have $\lambda_1 x = Ax = \lambda_2 x$. It's possible only when $x = 0$.

(e) Yes. By the hypothesis, we know A is diagonalizable. Say $A = P^{-1}DP$ for some invertible matrix P and some diagonal matrix D. Thus we know that

$$Q^{-1}AQ = (PQ)^{-1}D(PQ).$$

(f) No. It need one more condition that the characteristic polynomial spilts. For example, the matrix $\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$ has no real eigenvalue.

(g) Yes. Since it's a diagonalizable operator on nonzero vector space, it's characteristic polynomial spilts with degree greater than or equal to 1. So it has at least one zero.

(h) Yes. Because we have

$$W_i \cup \sum_{i \neq k} W_k = \{0\}$$

and

$$\sum_{i \neq k} W_k = \{0\} \supset W_j$$

for all $j \neq i$, we get the desired answer.

(i) No. For example, take $W_1 = \text{span}\{(1,0)\}$, $W_2 = \text{span}\{(0,1)\}$, and $W_3 = \text{span}\{(1,1)\}$.

2. For these question, see the direction of the subsection "Test for Diagonalization".

   (a) It's not diagonalizable since $\dim(E_1)$ is 1 but not 2.

   (b) It's diagonalizable with $D = \begin{pmatrix} -2 & 0 \\ 0 & 4 \end{pmatrix}$ and $Q = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$.

   (c) It's diagonalizable with $D = \begin{pmatrix} -2 & 0 \\ 0 & 5 \end{pmatrix}$ and $Q = \begin{pmatrix} 4 & 1 \\ -3 & 1 \end{pmatrix}$.

   (d) It's diagonalizable with $D = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ and $Q = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 0 & 4 \\ 0 & 1 & 3 \end{pmatrix}$.

   (e) It's not diagonalizable since its characteristic polynomial does not split.

   (f) It's not diagonalizable since $\dim(E_1)$ is 1 but not 2.

   (g) It's diagonalizable with $D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ and $Q = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & 1 \\ -1 & -1 & -1 \end{pmatrix}$.

3. For these question, we may choose arbitrary matrix representation, usually use the standard basis, and do the same as what we did in the previous exercises. So here we'll have $[T]_\beta = D$ and the set of column vectors of Q is the ordered basis β.

   (a) It's not diagonalizable since $\dim(E_0)$ is 1 but not 4.

   (b) It's diagonalizable with $D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $Q = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}$.

   (c) It's not diagonalizable since its characteristic polynomial does not split.

   (d) It's diagonalizable with $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and $Q = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ -1 & 0 & 0 \end{pmatrix}$.

   (e) It's diagonalizable with $D = \begin{pmatrix} 1-i & 0 \\ 0 & i+1 \end{pmatrix}$ and $Q = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$.

   (f) It's diagonalizable with $D = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ and $Q = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

4. It's not funny again. Replace the character T by the character A in that prove.

5. It's not funny again and again. Replace the character T by the character A in that prove.

6. (a) An operator T is diagonalizable ensure that its characteristic polynomial splits by Theorem 5.6. And in this situation Theorem 5.9(a) ensure that the multiplicity of each eigenvalue meets the dimension of the corresponding eigenspace. Conversly, if the characteristic polynomial splits and the multiplicity meets the dimension, then the operator will be diagonalizable by Theorem 5.9(a).

   (b) Replace T by A again.

7. Diagonalize the matrix A by $Q^{-1}AQ = D$ with $D = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$ and $Q = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$. So we know that

$$A^n = QD^nQ^{-1} = Q\begin{pmatrix} 5^n & 0 \\ 0 & (-1)^n \end{pmatrix}Q^{-1}.$$

8. We know that $\dim(E_{\lambda_2}) \geq 1$. So pick a nonzero vector $v \in E_{\lambda_2}$. Also pick a basis β for E_{λ_2}. Then $\alpha = \beta \cup \{v\}$ forms a basis consisting of eigenvectors. It's a basis because the cardinality is n and the help of Theorem 5.8.

9. (a) Because the characteristic polynomial of T is independent of the choice of β, we know that the characteristic polynomial

$$f(t) = \det([T]_\beta - tI) = \prod_{i=1}^n \left(([T]_\beta)_{ii} - t\right)$$

   splits, where the second equality holds since it's a upper triangular matrix.

   (b) The characteristic polynomial of a matrix is also the same for all matrices which is similar to the original matrix.

10. This is because the equality in Exercise 5.2.9(a). That is, if $[T]_\beta$ is an upper triangular matrix, it's diagonal entries are the set of zeroes of the characteristic polynomial.

11. (a) Since eigenvalues are the zeroes of the characteristic polynomial, we may write the characteristic polynomial $f(t)$ as

$$(\lambda_1 - t)^{m_1}(\lambda_2 - t)^{m_2}\cdots(\lambda_k - t)^{m_k}.$$

   Calculate the coefficient of t^{n-1} and use the fact in Exercise 5.1.21(b). Thus we could get the desired conclusion.

   (b) Use the equality in the previous exercise and calculate the coefficient of the constant term. Compare it to the Exercise 5.1.20.

12. (a) Let E_λ be the eigenspace of T corresponding to λ and $E_{\lambda^{-1}}$ be the eigenspace of T^{-1} corresponding to λ^{-1}. We want to prove the two spaces are the same. If $v \in E_\lambda$, we have $T(v) = \lambda v$ and so $v = \lambda T^{-1}(v)$. This means $T^{-1}(v) = \lambda^{-1}v$ and $v \in E_{\lambda^{-1}}$. Conversely, if $v \in E_{\lambda^{-1}}$, we have $T^{-1}(v) = \lambda^{-1}v$ and so $v = \lambda^{-1}T(v)$. This means $T(v) = \lambda v$ and $v \in E_\lambda$.

(b) By the result of the previous exercise, if T is diagonalizable and invertible, the basis consisting of eigenvectors of T will also be the basis consisting of eigenvectors of T^{-1}.

13. (a) For matrix $A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$, corresponding to the same eigenvalue 0 we have $E_0 = \text{span}\{(0,1)\}$ is the eigenspace for A while $E_0 = \text{span}\{(1,-1)\}$ is the eigenspace for A^t.

(b) Observe that

$$\dim(E_\lambda) = \text{null}(A - \lambda I) = \text{null}((A - \lambda I)^t)$$
$$= \text{null}(A^t - \lambda I) = \dim(E'_\lambda).$$

(c) If A is diagonalizable, then its characteristic polynomial splits and the multiplicity meets the dimension of the corresponding eigenspace. Since A and A^t has the same characteristic polynomial, the characteristic polynomial of A^t also splits. And by the precious exercise we know that the multiplicity meets the dimension of the corresponding eigenspace in the case of A^t.

14. (a) Let $v = (x,y)$ and $v' = (x',y')$ and $A = \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix}$. We may write the system of equation as $Av = v'$. We may also diagonalize the matrix A by $Q^{-1}AQ = D$ with $D = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}$ and $Q = \begin{pmatrix} 1 & 1 \\ -3 & 1 \end{pmatrix}$. This means $D(Q^{-1}v) = (Q^{-1}v)'$. So we know that

$$Q^{-1}v = \begin{pmatrix} c_1 e^{-2t} \\ c_2 e^{2t} \end{pmatrix}$$

and

$$v = Q\begin{pmatrix} c_1 e^{-2t} \\ c_2 e^{2t} \end{pmatrix},$$

where c_i is some scalar for all i.

(b) Calculate $D = \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}$ and $Q = \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix}$. So we have

$$v = Q\begin{pmatrix} c_1 e^{3t} \\ c_2 e^{-2t} \end{pmatrix},$$

where c_i is some scalar for all i.

(c) Calculate $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ and $Q = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. So we have

$$v = Q \begin{pmatrix} c_1 e^t \\ c_2 e^t \\ c_3 e^{2t} \end{pmatrix},$$

where c_i is some scalar for all i.

15. Following the step of the previous exercise, we may pick a matrix Q whose column vectors consist of eigenvectors and Q is invertible. Let D be the diagonal matrix $Q^{-1}AQ$. And we also know that finally we'll have the solution $x = Qu$ for some vector u whose i-th entry is $c_i e^\lambda$ if the i-th column of Q is an eigenvector corresponding to λ. By denoting $\bar{D}$ to be the diagonal matrix with $\bar{D}_{ii} = e^{D_{ii}}$, we may write $x = Q\bar{D}y$. where the i-th entry of y is c_i. So the solution must be of the form described in the exercise.

For the second statement, we should know first that the set

$$\{e^{\lambda_1 t}, e^{\lambda_2 t}, \dots e^{\lambda_k t}\}$$

are linearly independent in the space of real functions. Since Q invertible, we know that the solution set

$$\{Q\bar{D}y : y \in \mathbb{R}^n\}$$

is an n-dimensional real vector space.

16. Directly calculate that

$$(CY)'_{ij} = \left(\sum_{k=1}^{n} C_{ik}Y_{kj} \right)'$$

$$= \sum_{k=1}^{n} C_{ik}Y'_{kj} = CY'.$$

17. (a) We may pick one basis α such that both $[T]_\alpha$ and $[U]_\alpha$ are diagonal. Let $Q = [I]_\alpha^\beta$. And we will find out that

$$[T]_\alpha = Q^{-1}[T]_\beta Q$$

and

$$[U]_\alpha = Q^{-1}[U]_\beta Q.$$

(b) Let Q be the invertible matrix who makes A and B simultaneously diagonalizable. Say β be the basis consisting of the column vectors of Q. And let α be the standard basis. Now we know that

$$[T]_\beta = [I]_\alpha^\beta [T]_\alpha [I]_\beta^\alpha = Q^{-1}AQ$$

and

$$[U]_\beta = [I]_\alpha^\beta [U]_\alpha [I]_\beta^\alpha = Q^{-1}BQ.$$

18. (a) Let β be the basis makes T and U simultaneously diagonalizable. We know that each pair of diagonal matrices commute. So we have

$$[T]_\beta[U]_\beta = [U]_\beta[T]_\beta.$$

And this means T and U commute.

(b) Let Q be the invertible matrix who makes A and B simultaneously diagonalizable. Thus we have

$$(Q^{-1}AQ)(Q^{-1}BQ) = (Q^{-1}BQ)(Q^{-1}AQ).$$

And this means that A and B commute since Q is invertible.

19. They have the same eigenvectors by Exercise 5.1.15.

20. Since we can check that

$$W_1 \oplus W_2 \oplus W_3 = (W_1 \oplus W_2) \oplus W_3,$$

if V is direct sum of $W_1, W_2, \dots W_k$, the dimension of V would be the sum of the dimension of each W_i by using Exercise 1.6.29(a) inductively.

Conversely, we first prove that if we have

$$\sum_{i=1}^{k} W_i = V,$$

then we must have

$$\dim(V) \le \sum_{i=1}^{k} \dim(W_i).$$

We induct on k. For $k = 2$, we may use the formula in Exercise 1.6.29(a). Suppose it holds for $k = m$. We know that

$$\left(\sum_{i=1}^{m-1} W_i \right) + W_m = \sum_{i=1}^{m} W_i$$

and

$$\dim(\sum_{i=1}^{m} W_i) \le \dim(\sum_{i=1}^{m-1} W_i) + \dim(W_m) \le \sum_{i=1}^{m} \dim(W_i)$$

by induction hypothesis and the case for $k = 2$.

To prove the original statement, suppoe, by contradiction, that

$$W_1 \cap \sum i = 2^k W_i$$

has nonzero element. By the formula in Exercise 1.6.29(a) we know that

$$\dim(\sum i = 2^k W_i) > \dim(V) - \dim(W_1) = \sum_{i=2}^{k} \dim(W_i).$$

This is impossible, so we get the desired result.

21. Because β is a basis, we have

$$\sum_{i=1}^{k} \text{span}(\beta_i) = V.$$

Second, since the dimension of $\text{span}(\beta_i)$ is the number of element of β_i, we know that the equality in the previous exercise about dimension holds. So V is the direct sum of them.

22. By the definition of the left hand side, it is the sum of each eigenspaces. Let $W = \sum_{i=2}^{k} E_{\lambda_i}$. If there is some nonzero vector v_1 in $E_{\lambda_1} \cap W$. We may write

$$v_1 + c_2 v_2 + c_3 v_3 + \cdots + c_k v_k = 0$$

for some scalar c_i and some eigenvectors $v_i \in E_{\lambda_i}$. Now we know that

$$0 = T(0) = \lambda_1 v_1 + c_2 \lambda_2 v_2 + c_3 \lambda_3 v_3 + \cdots + c_k \lambda_k v_k = 0.$$

After substracting this equality by λ_1 times the previous equality, we get

$$c_2(\lambda_2 - \lambda_1)v_2 + \cdots + c_k(\lambda_k - \lambda_1)v_k = 0.$$

This is impossible since $\lambda_i - \lambda_1$ is nonzero for all i and c_i cannot be all zero. Similarly we know that E_{λ_i} has no common element other than zero with the summation of other eigenspaces. So the left hand side is the direct sum of each eigenspaces.

23. It's enough to check whether K_1 has common nonzero element with the summation of others or not. Thus we can do similarly for the other cases. Let V_1 be the summation of $K_2, K_3, \ldots, K_p$. Now, if there is some nonzero vector $x \in K_1(V_1 \cap W_2)$, we may assume that $x = u + v$ with $u \in V_1$ and $v \in W_2$. Since W_1 is the direct sum of all K_i's, we know $K_1 \cap V_1 = \{0\}$. So $x - u = v \neq 0$ is an element in both W_1 and W_2, a contradiction. Thus we've completed the proof.

5.3 Matrix Limits and Markov Chains

1. (a) Yes. This is the result of Theorem 5.12.

   (b) Yes. This is the result of Theorem 5.13.

   (c) No. It still need the condition that each entry is nonnegative.

   (d) No. It's the sum of each column. The matrix $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ is a counterexample.

   (e) Yes. See the Corollary after Theorem 5.15, although there's no proof.

   (f) Yes. See the Corollary 1 after Theorem 5.16.

   (g) Yes. Look Theorem 5.17.

(h) No. The matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has eigenvector $(1, -1)$ corresponding the eigenvalue -1.

(i) No. The matrix $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has the property that $A^2 = I$. So the sequence would be $A, I, A, I, \ldots$, which does not converge.

(j) Yes. This is Theorem 5.20.

2. Diagonalize those matrices and use the fact of the Corollary after Theorem 5.12. Actually the eigenvalue will not tend to zero if and only if the eigenvalue is 1.

(a) The limit is the zero matrix.

(b) The limit is $\begin{pmatrix} -0.5 & 0.5 \\ -1.5 & 1.5 \end{pmatrix}$.

(c) The limit is $\begin{pmatrix} \frac{7}{13} & \frac{7}{13} \\ \frac{6}{13} & \frac{6}{13} \end{pmatrix}$.

(d) The limit is the zero matrix.

(e) The limit does not exist.

(f) The limit is $\begin{pmatrix} 3 & -1 \\ 6 & -2 \end{pmatrix}$.

(g) The limit is $\begin{pmatrix} -1 & 0 & -1 \\ -4 & 1 & -2 \\ 2 & 0 & 2 \end{pmatrix}$.

(h) The limit is $\begin{pmatrix} -2 & -3 & -1 \\ 0 & 0 & 0 \\ 6 & 9 & 3 \end{pmatrix}$.

(i) The limit does not exist.

(j) The limit does not exist.

3. We know that

$$\lim_{m \to \infty} (A_m)^t_{ij} = \lim_{m \to \infty} (A_m)_{ji} = L_{ji} = L^t_{ij}.$$

4. If A is diagonalizable, we may say that $Q^{-1}AQ = D$ for some invertible matrix Q and for some diagonal matrix D whose diagonal entries are eigenvalues. So $\lim_{m \to \infty} D^m$ exist only when all its eigenvalues are numbers in S, which was defined in the paragraphs before Theorem 5.13. If all the eigenvalues are 1, then the limit L would be I_n. If some eigenvalue $\lambda \neq 1$, then its absolute value must be less than 1 and the limit of λ^m would shrink to zero. This means that L has rank less than n.

5. First we see three special matrices,
$$P = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, Q = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, R = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

The limit of power of these three matrices is themselves respectly, because they has the property that their square are themselves. However, we know that
$$\frac{1}{\sqrt{2}}P \cdot \frac{1}{\sqrt{2}}P = R.$$

So we pick $A = \frac{1}{\sqrt{2}}P$, $B = \frac{1}{\sqrt{2}}Q$, and $C = R$. Thus we have the limit of power of A and B are the zero matrix. And the limit of power of C is C itself.

6. Let x, y, z, and w denote the percentage of the healthy, ambulatory, bedridden, and dead patients. And we know that
$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0.3 \\ 0.7 \\ 0 \end{pmatrix}$$

and the percentage of each type would be described by the vector
$$\begin{pmatrix} 1 & 0.6 & 0.1 & 0 \\ 0 & 0.2 & 0.2 & 0 \\ 0 & 0.2 & 0.5 & 0 \\ 0 & 0 & 0.2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0.3 \\ 0.7 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.25 \\ 0.2 \\ 0.41 \\ 0.14 \end{pmatrix}$$

in the same order. So this answer the first question. For the second question we should calculate the limit of power of that transition matrix, say A. It would be
$$L = \lim_{m \to \infty} A^m = \begin{pmatrix} 1 & \frac{8}{9} & \frac{5}{9} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \frac{1}{9} & \frac{4}{9} & 1 \end{pmatrix}.$$

So the limit of percentage of each type would be described by the vector
$$L \begin{pmatrix} 0 \\ 0.3 \\ 0.7 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{59}{90} \\ 0 \\ 0 \\ \frac{31}{90} \end{pmatrix}.$$

7. Using the language of the stochastic matrix, we direct write the stochastic matrix to be
$$A = \begin{pmatrix} 1 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 \\ 0 & \frac{2}{3} & 0 & 0 \\ 0 & 0 & \frac{2}{3} & 1 \end{pmatrix}.$$

I don't think there is too much difference between the process of diagonalizing and finding the eigenvectors. It's much easier to observe the sequence

$$e_2, Ae_2, A^2 e_2, \ldots,$$

which is

$$\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{3} \\ 0 \\ \frac{2}{3} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{3} \\ \frac{2}{9} \\ 0 \\ \frac{4}{9} \end{pmatrix}, \begin{pmatrix} \frac{11}{27} \\ 0 \\ \frac{4}{27} \\ \frac{4}{9} \end{pmatrix}, \begin{pmatrix} \frac{11}{27} \\ \frac{4}{81} \\ 0 \\ \frac{44}{81} \end{pmatrix}, \ldots$$

And find the limit of the first entry to be

$$\frac{1}{3} + \frac{2}{3} \cdot \frac{1}{9} + \frac{2}{3} \cdot \frac{2}{9} \cdot \frac{1}{9} + \frac{2}{3} \cdot \left(\frac{2}{9}\right)^2 \cdot \frac{1}{9} + \cdots$$

$$\frac{1}{3} + \frac{\frac{2}{3} \cdot \frac{1}{9}}{1 - \frac{2}{9}} = \frac{3}{7}.$$

So the answer is $\frac{3}{7}$.

8. There's no better method to check whether the matrix is regular or not. So just try it or find the evidence that the power of it will not be a positive matrix. When the matrix is nonnegative, we may just consider the matrix obtained by replacing each nonzero entry by 1.

   (a) Yes. The square of it is positive.
   (b) Yes. It's positive when the power is 4.
   (c) No. The second and the third column always interchange each time.
   (d) No. The second column does not change each time.
   (e) No. The second and the third columns do not change.
   (f) No. The first column does not change.
   (g) No. The third and the fourth columns do not change.
   (h) No. The third and the fourth columns do not change.

9. Use the same method as that in Exercise 5.3.2. Or we may use the result of 5.20 for the case of regular matrix.

   (a) The limit is $\begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$.

   (b) The limit is $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$.

   (c) The limit does not exist since the second and the third columns interchange each time.

(d) The limit is $\begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

(e) The limit is $\begin{pmatrix} 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{pmatrix}$.

(f) The limit is $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{2}{5} & \frac{2}{5} \\ 0 & \frac{3}{5} & \frac{3}{5} \end{pmatrix}$.

(g) The limit is $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 1 \end{pmatrix}$.

(h) The limit is $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 1 \end{pmatrix}$.

10. To calculate the vector after two stage, we could only multiply the matrix twice and multiply the vectors. To calculate the limit matrix we could just find the eigenvector corresponding to eigenvalue 1 and use Theorem 5.20.

(a) The two-stage vectors and the fixed vectors are

$$\begin{pmatrix} 0.225 \\ 0.441 \\ 0.334 \end{pmatrix}, \begin{pmatrix} 0.2 \\ 0.6 \\ 0.2 \end{pmatrix}$$

respectively.

(b) The two-stage vectors and the fixed vectors are

$$\begin{pmatrix} 0.375 \\ 0.375 \\ 0.25 \end{pmatrix}, \begin{pmatrix} 0.4 \\ 0.4 \\ 0.2 \end{pmatrix}$$

respectively.

(c) The two-stage vectors and the fixed vectors are

$$\begin{pmatrix} 0.372 \\ 0.225 \\ 0.403 \end{pmatrix}, \begin{pmatrix} 0.5 \\ 0.2 \\ 0.3 \end{pmatrix}$$

respectively.

(d) The two-stage vectors and the fixed vectors are

$$\begin{pmatrix} 0.252 \\ 0.334 \\ 0.414 \end{pmatrix}, \begin{pmatrix} 0.25 \\ 0.35 \\ 0.4 \end{pmatrix}$$

respectly.

(e) The two-stage vectors and the fixed vectors are

$$\begin{pmatrix} 0.329 \\ 0.334 \\ 0.337 \end{pmatrix}, \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$$

respectly.

(f) The two-stage vectors and the fixed vectors are

$$\begin{pmatrix} 0.316 \\ 0.428 \\ 0.256 \end{pmatrix}, \begin{pmatrix} 0.25 \\ 0.5 \\ 0.25 \end{pmatrix}$$

respectly.

11. The matrix would be

$$A = \begin{pmatrix} 0.7 & 0.2 & 0 \\ 0.1 & 0.6 & 0.2 \\ 0.2 & 0.2 & 0.8 \end{pmatrix}.$$

So the vector in 1950 would be

$$A \begin{pmatrix} 0.1 \\ 0.5 \\ 0.4 \end{pmatrix} = \begin{pmatrix} 0.197 \\ 0.339 \\ 0.464 \end{pmatrix}.$$

Since it's regular, we may just find the fixed vector

$$\begin{pmatrix} 0.2 \\ 0.3 \\ 0.5 \end{pmatrix}.$$

12. Considering the three states, new, once used, and twice used, the matrix is

$$A = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & 1 \\ \frac{2}{3} & 0 & 0 \\ 0 & \frac{2}{3} & 0 \end{pmatrix}.$$

It is regular. So we can just find the fixed vector

$$\begin{pmatrix} \frac{9}{19} \\ \frac{6}{19} \\ \frac{4}{19} \end{pmatrix}.$$

13. Considering the three states, large, intermediate-sized, and small car owners, the matrix is

$$\begin{pmatrix} 0.7 & 0.1 & 0 \\ 0.3 & 0.7 & 0.1 \\ 0 & 0.2 & 0.9 \end{pmatrix}$$

and the initial vector is

$$P = \begin{pmatrix} 0.4 \\ 0.2 \\ 0.4 \end{pmatrix}.$$

So the vector in 1995 would be

$$A^2 P = \begin{pmatrix} 0.24 \\ 0.34 \\ 0.42 \end{pmatrix}.$$

And the matrix A is regular. So we may just find the fixed vector

$$\begin{pmatrix} 0.1 \\ 0.3 \\ 0.6 \end{pmatrix}.$$

14. We prove it by induction on m. When $m = 1$, the formula meet the matrix A. Suppose the formula holds for the case $m = k - 1$. Then the case $m = k$ would be

$$A^k = A^{k-1} A = \begin{pmatrix} r_{k-1} & r_k & r_k \\ r_k & r_{k-1} & r_k \\ r_k & r_k & r_{k-1} \end{pmatrix} A$$

$$= \begin{pmatrix} r_k & \frac{r_{k-1}+r_k}{2} & \frac{r_{k-1}+r_k}{2} \\ \frac{r_{k-1}+r_k}{2} & r_k & \frac{r_{k-1}+r_k}{2} \\ \frac{r_{k-1}+r_k}{2} & \frac{r_{k-1}+r_k}{2} & r_k \end{pmatrix}.$$

After check that

$$\frac{r_{k-1} + r_k}{2} = \frac{1}{3} \cdot \frac{1 + \frac{(-1)^k}{2^{k-2}} + 1 + \frac{(-1)^k}{2^{k-1}}}{2} = \frac{1}{3}\left[1 + \frac{(-1)^k}{2^k}\right] = r_{k+1}$$

we get the desired result. To deduce the second equality, just replace A^m by the right hand side in the formula.

15. Let v be that nonnegative vector and say d is the sum of all its entries. Thus we have $x = \frac{1}{d} v$ is a probability vector. Furthermore, if y is a probability vector in W. They must be parallel, say $y = kx$. The sum of entries of y is k by the fact that x is a probability vector. And the sum of entries of y is 1 since itself is a probability vector. This means $k = 1$. So the vector is unique.

16. If A is a (not necessarily square) matrix with row vectors $A_1, A_2, \ldots, A_r$, we have the fact
$$A^t u = A_1^t + A_2^t + \cdots + A_r^t.$$

This vector equal to u if and only if the sum of entries is 1 for every column of A. Use this fact we've done the proof of Theorem 5.15.

For the Corollary after it, if M is a transition matrix, we have

$$(M^k)^t u = (M^t)^k u = (M^t)^{k-1} u = \cdots = u.$$

And if v is probability vector,

$$(Mv)^t u = v^t M^t u = v^t u = \left(1\right).$$

By Theorem 5.15 we get the conclusion.

17. For the first Corollary, we may apply Theorem 5.18 to the matrix A^t since we have the fact that A and A^t have the same characteristic polynomial. Thus we know that $\lambda = \nu(A)$. Also, the dimension of eigenspace of A^t corresponding to λ is 1. But Exercise 5.2.13 tell us that A^t and A has the same dimension of the corresponding eigenspaces.

For the second Corollary, we know that $\nu(A) = 1$. So if $\lambda \neq 1$ then we have $|\lambda| < 1$ by Theorem 5.18 and its first Corollary. And the eigenspace corresponding 1 has dimension one by the first Corollary.

18. By Theorem 5.19, all eigenvalues lies in S, which was defined in the paragraphs before Theorem 5.13. Thus the limit exist by Theorem 5.14.

19. (a) First we check that

$$(cM + (1-c)N)^t u = cM^t u + (1-c)N^t u = cu + (1-c)u = u.$$

Thus the new matrix is a transition matrix by the Corollary after Theorem 5.15. So it's enough to show that the new matrix is regular. Now suppose that M^k is positive. Then we know that $(cM + (1-c)N)^k$ is the sum of $c^k M^k$ and some lower order terms, which are nonnegative. So we know that it's a positive matrix and so the new matrix is regular.

(b) Pick a scalar d such that each entry of dM' is larger than that of M. Then we may pick $c = \frac{1}{d}$ and

$$N = \frac{1}{1-c}(M' - cM)$$

and know that N is nonnegative. Finally we may check that

$$N^t u = \frac{1}{1-c}(M'^t u - cM^t u) = \frac{1}{1-c} \cdot (1-c)u = u.$$

So N is also a transition matrix by the Corollary after Theorem 5.15.

(c) By symmetry, it's enough to prove only the one side of that statement. Also, by (b) we could write

$$M' = cM + (1 - c)N$$

for some scalar c and some transition matrix N. Now, if M is regular, then M' is also regular by (a).

20. Use the notation of the Definition, when $A = O$ we can observe that B_m is equal to I for all m. So e^O would be I. For the case $A = I$, we have $A^m = I$ for all m. So the matrix $e^I = eI$. That is, a matrix with all its entries e.

21. We set

$$B_m = I + A + \frac{A^2}{2!} + \cdots + \frac{A^m}{m!}$$

and

$$E_m = I + D + \frac{D^2}{2!} + \cdots + \frac{A^m}{m!}.$$

We may observe that $B_m = PE_m P^{-1}$ for all m. So by the definition of exponential of matrix and Theorem 5.12, we have

$$e^A = \lim_{m \to \infty} B_m = \lim_{m \to \infty} (PE_m P^{-1}) = P(\lim_{m \to \infty} E_m)P^{-1} = Pe^D P^{-1}.$$

22. With the help of previous exercise, it's enough to show that e^D exist if $P^{-1}AP = D$. Since we know that

$$(D^m)_{ii} = (D_{ii})^m,$$

we have that ,with the notation the same as the previous exercise,

$$(E_m)_{ii} = 1 + D_{ii} + \frac{(D_{ii})^2}{2!} + \cdots + \frac{(D_{ii})^m}{m!},$$

which tends to $e^{D_{ii}}$ as m tends to infinity. So we know that e^D exists.

23. The equality is usually not true, although it's true when A and B are diagonal matrices. For example, we may pick $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. By some calculation we have

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

and

$$e^B = \begin{pmatrix} e & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$e^A e^B = \begin{pmatrix} e^2 & e-1 \\ 0 & 1 \end{pmatrix}.$$

However, we may also calculate that

$$e^{A+B} = \begin{pmatrix} e^2 & \frac{e^2}{2} - \frac{1}{2} \\ 0 & 1 \end{pmatrix}.$$

24. Use the notation in the solution to Exercise 5.2.15. And in the proof we know that the solution $x(t) = Q\bar{D}y$ for some $y \in \mathbb{R}^n$. Now we know that actually $\bar{D} = e^D$ by definition. With the fact that Q is invertible, we may write the set of solution to be

$$\{Qe^D Q^{-1} v : v \in \mathbb{R}^n\} = \{e^A v : v \in \mathbb{R}^n\}$$

by Exercise 5.3.21.

5.4 Invariant Subspace and the Cayley-Hamilton Theorem

1. (a) No. The subspace $\{0\}$ must be a T-invariant subspace.

   (b) Yes. This is Theorem 5.21.

   (c) No. For example, let T be the identity map from $\mathbb{R}$ to $\mathbb{R}$ and $v = (1)$ and $w = (2)$. Then we have $W = W' = \mathbb{R}$.

   (d) No. For example, let T be the mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$ defined by $T(x, y) = y$. Pick $v = (1, 1)$. Thus the T-cyclic subspace generated by v and $T(v)$ are $\mathbb{R}^2$ and the x-axis.

   (e) Yes. The characteristic polynomial is the described polynomial.

   (f) Yes. We may prove it by induction or just use Exercise 5.1.21.

   (g) Yes. This is Theorem 5.25.

2. (a) Yes. For every $ax^2 + bx + c \in W$, we have

   $$T(ax^2 + bx + c) = 2ax + b \in W.$$

   (b) Yes. For every $ax^2 + bx + c \in W$, we have

   $$T(ax^2 + bx + c) = 2ax^2 + bx \in W.$$

   (c) Yes. For every $(t, t, t) \in W$, we have

   $$T(t, t, t) = (3t, 3t, 3t) \in W.$$

(d) Yes. For every $at + b \in W$, we have

$$T(at + b) = (\frac{a}{2} + b)t \in W.$$

(e) No. For $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \in W$, we have

$$T(A) = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \notin W.$$

3. (a) We have $T(0) = 0$ and $T(v) \in V$ any $v \in V$ and for arbitrary linear operator T on V.

(b) If $v \in N(T)$, we have $T(v) = 0 \in N(T)$. If $v \in R(T)$, we have $T(v) \in R(T)$ by the definition of $R(T)$.

(c) If $v \in E_\lambda$, we have $T(v) = \lambda v$. Since $T(\lambda v) = \lambda T(v) = \lambda^2 v$, we know that $T(v) \in E_\lambda$.

4. We have
$$T^k(W) \subset T^{k-1}(T(W)) \subset T^{k-1}(W) \subset \cdots \subset W.$$

For any $v \in W$, we know that $g(T)(v)$ is the sum of some elements in W. Since W is a subspace, we know that $g(T)(v)$ is always an element is W.

5. Let $\{W_i\}_{i \in I}$ be the collection of T-invatirant subspaces and W be the intersection of them. For every $v \in W$, we have $T(v) \in W_i$ for every $i \in I$, since v is an element is each W_i. This means $T(v)$ is also an element in W.

6. Follow the prove of Theorem 5.22. And we know that the dimension is the maximum number k such that

$$\{z, T(z), T^2(z), \ldots, T^{k-1}(z)\}$$

is independent and the set is a basis of the subspace.

(a) Calculate that

$$z = (1, 0, 0, 0), T(z) = (1, 0, 1, 1),$$

$$T^2(z) = (1, -1, 2, 2), T^3(z) = (0, -3, 3, 3).$$

So we know the dimension is 3 and the set

$$\{z, T(z), T^2(z)\}$$

is a basis.

(b) Calculate that
$$z = x^3, T(z) = 6x, T^2(z) = 0.$$

So we know that the dimension is 2 and the set
$$\{z, T(z)\}$$

is a basis.

(c) Calculate that $T(z) = z$. So we know that the dimension is 1 and $\{z\}$ is a basis.

(d) Calculate that
$$z = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, T(z) = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix},$$
$$T^2(z) = \begin{pmatrix} 3 & 3 \\ 6 & 6 \end{pmatrix}.$$

So we know that the dimension is 2 and the set
$$\{z, T(z)\}$$

is a basis.

7. Let W be a T-invariant subspace and T_W be the restricted operator on W. We have that
$$R(T_W) = T_W(W) = T(W) \subset W.$$

So at least it's a well-defined mapping. And we also have
$$T_W(x) + T_W(y) = T(x) + T(y) = T(x+y) = T_W(x+y)$$

and
$$T_W(cx) = T(cx) = cT(x) = cT_W(x).$$

So the restriction of T on W is also a lineaer operator.

8. If v is an eigenvector of T_W corresponding eigenvalue λ, this means that $T(v) = T_W(v) = \lambda v$. So the same is true for T.

9. See Example 5.4.6.

(a) For the first method, we may calculate $T^3(z) = (3, -3, 3, 3)$ and represent it as a linear combination of the basis
$$T^3(z) = 0z - 3T(z) + 3T^2(z).$$

So the characteristic polynomial is $-t^3 + 3t^2 - 3t$. For the second method, denote β to be the ordered basis
$$\{z, T(z), T^2(z), T^3(z)\}.$$

And we may calculate the matrix representation

$$[T_W]_\beta = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -3 \\ 0 & 1 & 3 \end{pmatrix}$$

and directly find the characteristic polynomial of it to get the same result.

(b) For the first method, we may calculate $T^3(z) = 0$ and represent it as a linear combination of the basis

$$T^2(z) = 0z + 0T(z).$$

So the characteristic polynomial is t^2. For the second method, denote β to be the ordered basis

$$\{z, T(z)\}.$$

And we may calculate the matrix representation

$$[T_W]_\beta = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

and directly find the characteristic polynomial of it to get the same result.

(c) For the first method, we may calculate $T(z) = z$. So the characteristic polynomial is $-t + 1$. For the second method, denote β to be the ordered basis $\{z\}$. And we may calculate the matrix representation

$$[T_W]_\beta = \begin{pmatrix} 1 \end{pmatrix}$$

and directly find the characteristic polynomial of it to get the same result.

(d) For the first method, we may calculate $T^2(z) = 3T(z)$. So the characteristic polynomial is $t^2 - 3t$. For the second method, denote β to be the ordered basis
$$\{z, T(z)\}.$$

And we may calculate the matrix representation

$$[T_W]_\beta = \begin{pmatrix} 0 & 0 \\ 1 & 3 \end{pmatrix}$$

and directly find the characteristic polynomial of it to get the same result.

10. Calculate the characteristic polynomial is the problem of the first section and determine whether on polynomial is divided by the other is the problem in senior high.

(a) The characteristic polynomial is

$$t^4 - 4t^3 + 6t^2 - 3t.$$

(b) The characteristic polynomial is t^4.

(c) The characteristic polynomial is

$$t^4 - 2t^3 + 2t - 1.$$

(d) The characteristic polynomial is

$$t^4 - 6t^3 + 9t^2.$$

11. (a) Let w be an element in W. We may express w to be

$$w = \sum_{i=0}^{k} a_i T^i(v).$$

And thus we have

$$T(w) = \sum_{i=0}^{k} a_i T^{i+1}(v) \in W.$$

since U contains all of T^k(v), then U also must contains span of {v, T(v), ...}, which is the T-invariant subspace W

(b) Let U be a T-invariant subspace of V containing v. Since it's T-invariant, we know that $T(v)$ is an element in U. Inductively, we know that $T^k(v) \in U$ for all nonnegative integer k. By Theorem 1.5 we know that U must contain W.

12. Because W is a T-invariant subspace, we know that $T(v_i) \in W$ for $v_i \in \gamma$. This means the representation of $T(v_i)$ corresponding the basis β only use the vectors in γ for each $v_i \in \gamma$. So one corner of the matrix representation would be zero.

13. If w is an element in W, it's a linear combination of

$$\{v, T(v), T^2(v), \ldots.$$

So $w = g(T)(v)$ for some polynomial g. Conversely, if $w = g(T)(v)$ for some polynomial g, this means w is a linear combination of the same set. Hence w is an element in W.

14. It's because that the set

$$v, T(v), T^2(v), \ldots, T^{k-1}(v)$$

is a basis of W by Theorem 5.22, where k is the dimension of W.

15. The question in the Warning is because the definition of $f(A)$ is not $\det(A - AI)$. To prove the version for matrix, we may apply the theorem to the linear transformation L_A. So we know that the characteristic polynomial of L_A is the same as that of A, say $f(t)$. And we get the result $f(L_A) = T_0$ by the theorem. This means

$$f(A)v = f(L_A)(v) = 0$$

for all v. So we know $f(A) = O$.

16. (a) By Theorem 5.21 we know that the characteristic polynomial of the restriction of T to any T-invariant subspace is a factor of a polynomial who splits. So it splits, too.

 (b) Any nontrivial T-invariant subspace has dimension not equal to 0. So the characteristic polynomial of its restriction has degree greater than or equal to 1. So it must contains at least one zero. This means the subspace at least contains one eigenvector.

17. If we have the characteristic polynomial to be

$$f(t) = (-1)^n t^n + a_{n-1} t^{n-1} + \cdots + a_0,$$

then we have

$$f(A) = (-1)^n A^n + a_{n-1} A^{n-1} + \cdots + a_0 I = O$$

by Cayley-Hamilton Theorem. This means that A^n is a linear combination of $I, A, A^2, \ldots, A^{n-1}$. By multiplying both sides by A, we know that A^{n+1} is a linear combination of $A, A^2, \ldots, A^n$. Since A^n can be represented as a linear combination of previous terms, we know that A^{n+1} could also be a linear combination of $I, A, A^2, \ldots, A^{n-1}$. Inductively, we know that

$$\text{span}\{I, A, A^2, \ldots\} = \text{span}\{I, A, A^2, \ldots, A^{n-1}\}$$

and so the dimension could not be greater than n.

18. (a) See Exercise 5.1.20.

 (b) For simplicity, denote the right hand side to be B. Directly calculate that
$$AB = -\frac{1}{a_0}[(-1)^n A^n + a_{n-1} A^{n-1} + \cdots + a_1 A]$$
$$-\frac{1}{a_0}(-a_0 I) = I.$$

 (c) Calculate that the characteristic polynomial of A is
$$-t^3 + 2t^2 + t - 2.$$

137

So by the previous result, we know that

$$A^{-1} = \frac{1}{2}[-A^2 + 2A + I$$

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

19. As Hint, we induct on k. For $k = 1$, the matrix is $\begin{pmatrix} -a_0 \end{pmatrix}$ and the characteristic polynomial is $-(a_0 + t)$. If the hypothesis holds for the case $k = m - 1$, we may the expand the matrix by the first row and calculate the characteristic polynomial to be

$$\det(A - tI) = \det \begin{pmatrix} -t & 0 & \cdots & -a_0 \\ 1 & -t & \cdots & -a_1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -a_{m-1} \end{pmatrix}$$

$$= -t \det \begin{pmatrix} -t & 0 & \cdots & -a_0 \\ 1 & -t & \cdots & -a_1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -a_{m-2} \end{pmatrix} + (-1)^{m+1}(-a_0) \det \begin{pmatrix} 1 & -t & \cdots & 0 \\ 0 & 1 & \cdots & -a_1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

$$= -t[(-1)^{m-1}(a_1 + a_2 t + \cdots + a_{m-1} t^{m-2} + t^{m-1})] + (-1)^m a_0$$

$$= (-1)^m (a_0 + a_1 t + \cdots + a_{m-1} t^{m-1} + t^m).$$

20. If $U = g(T)$, we know that $UT = TU$ since

$$T(T^k) = (T^k)T = T^{k+1}$$

and T is linear. For the converse, we may suppose that V is generated by v. Then the set

$$\beta = \{v, T(v), T^2(v), \ldots, T^k(v)\}$$

is a basis. So the vector $U(v)$ could be written as a linear combination of β. This means $U(v) = g(T)(v)$ for some polynomial g. Now if $UT = TU$, we want to show $U = g(T)$ by showing $U(w) = g(T)(w)$ for all $w \in \beta$. Observe that

$$U(T^m(v)) = T^m(U(v)) = T^m g(T)(v) = g(T)(T^m(v))$$

for all nonnegative integer m. So we get the desired result.

21. If we have some vector $v \in V$ such that v and $T(v)$ are not parallel, we know the T-cyclic subspace generated by v has dimension 2. This means V is a T-cyclic subspace of itself generated by v. Otherwise for every $v \in V$, we may find some scalar λ_v such that $T(v) = \lambda_v v$. Now, if λ_v is a same value c for every nonzero vector v, then we have $T = cI$. If not, we

may find $\lambda_v \neq \lambda_u$ for some nonzero vector v and u. This means v and u lies in different eigenspace and so the set $\{u, v\}$ is independent. Thus they forms a basis. Now let $w = v + u$. We have both

$$T(w) = \lambda_w w = \lambda_w v + \lambda_w u$$

and

$$T(w) = \lambda_v v + \lambda_u u.$$

By the uniqueness of representation of linear combination, we must have

$$\lambda_v = \lambda_w = \lambda_u,$$

a contradiction. So in this case we must have $T = cI$.

22. If $T \neq cI$, then V must be a T-cyclic subspace of itself by Exercise 5.4.21. So by Exercise 5.4.20, we get the desired result.

23. As Hint, we prove it by induction on k. For $k = 1$, it's a natural statement that if v_1 is in W then v_1 is in W. If the statement is true for $k = m - 1$, consider the case for $k = m$. If we have $u = v_1 + v_2 + \cdots + v_m$ is in W, a T-invariant subspace, then we have

$$T(u) = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_m v_m \in W,$$

where λ_i is the distinct eigenvalues. But we also have $\lambda_m u$ is in W. So the vector

$$T(u) - \lambda_m u = (\lambda_1 - \lambda_m)v_1 + (\lambda_2 - \lambda_m)v_2 + \cdots + (\lambda_{m-1} - \lambda_m)v_{m-1}$$

is al so in W. Since those eigenvalues are distinct, we have $\lambda_i - \lambda_m$ is not zero for all $i \neq m$. So we may apply the hypothesis to

$$(\lambda_1 - \lambda_m)v_1, (\lambda_2 - \lambda_m)v_2 + \cdots + (\lambda_{m-1} - \lambda_m)v_{m-1}$$

and get the result that $(\lambda_i - \lambda_m)v_i$ is in W and so v_i is in W for all $i \neq m$. Finally, we still have

$$v_m = u - v_1 - v_2 - \cdots - v_{m-1} \in W.$$

24. Let T be a operator on V and W be a nontrivial T-invariant subspace. Also let E_λ be the eigenspace of V corresponding to the eigenvalue λ. We set $W_\lambda = E_\lambda \cap W$ to be the eigenspace of T_W corresponding to the eigenvalue λ. We may find a basis β_λ for W_λ and try to show that $\beta = \cup_\lambda \beta_\lambda$ is a basis for W. The set β is linearly independent by Theorem 5.8. Since T is diagonalizable, every vector could be written as a linear combination of eigenvectors corresponding to distinct eigenvalues, so are those vectors in W. But by the previous exercise we know that those eigenvectors used to give a linear combination to elements in W must also in W. This means every elements in W is a linear combination of β. So β is a basis for W consisting of eigenvectors.

25. (a) Let E_λ be a eigenspace of T corresponding to the eigenvalue λ. For every $v \in E_\lambda$ we have

$$TU(v) = UT(v) = \lambda U(v).$$

This means that E_λ is an U-invariant subspace. Applying the previous exercise to each E_λ, we may find a basis β_λ for E_λ such that $[U_{E_\lambda}]_{\beta_\lambda}$ is diagonal. Take β to be the union of each β_λ and then both $[T]_\beta$ and $[U]_\beta$ are diagonal simultaneously.

(b) Let A and B are two $n \times n$ matrices. If $AB = BA$, then A and B are simultaneously diagonalizable. To prove this we may apply the version of linear transformation to L_A and L_B.

26. Let $\{v_1, v_2, \ldots, v_n$ be those eigenvectors corresponding to n distinct eigenvalues. Pick
$$v = v_1 + v_2 + \cdots + v_n$$
and say W to be the T-cyclic subspace generated by v. By Exercise 5.4.23 we know that W must contains all v_i's. But this means the dimension of W is n and the set
$$\{v, T(v), \ldots, T^{n-1}(v)\}$$
is a basis by Theorem 5.22.

~~27.~~ (a) If $v + W = v' + W$, we know that $v - v'$ is an element in W by Exercise 1.3.31. Since W is T-invariant, we have

$$T(v) - T(v') = T(v - v')$$

is in W. So we have

$$T(v) + W = T(v') + W$$

and this means
$$\overline{T}(v + W) = \overline{T}(v' + W).$$

(b) Just check that

$$\overline{T}((v + v') + W) = T(v + v') + W$$

$$= (T(v) + W) + (T(v') + W) = \overline{T}(v + W) + \overline{T}(v' + W)$$

and

$$\overline{T}(cv + W) = T(cv) + W = c(T(v) + W) = c\overline{T}(v + W).$$

(c) For each $v \in V$ we might see

$$\eta T(v) = T(v) + W = \overline{T}(v + W) = \overline{T}\eta(v).$$

140

28. We use the notation given in Hint. Since W is T-invariant, we know the matrix representation of T is

$$[T]_\beta = \begin{pmatrix} B_1 & B_2 \\ O & B_3 \end{pmatrix}.$$

As the proof of Theorem 5.21, we know that $f(t) = \det([T]_\beta - tI)$ and $g(t) = \det(B_1 - tI)$. It's enough to show $h(t) = \det(B_3 - tI)$ by showing B_3 is a matrix representation of $\overline{T}$. Let

$$\alpha = \{v_k + W, v_{k+1} + W, \ldots, v_n + W\}$$

be a basis for V/W by Exercise 1.6.35. Then for each $j = k, k+1, \ldots, n$, we have

$$\overline{T}(v_j) = T(v_j) + W$$

$$= [\sum_{l=1}^{k} (B_2)_{lj} v_l + \sum_{i=k+1}^{n} (B_3)_{ij} v_i] + W$$

$$= \sum_{i=k+1}^{n} (B_3)_{ij} v_i + W = \sum_{i=k+1}^{n} (B_3)_{ij} (v_i + W).$$

So we have $B_3 = [\overline{T}]_\alpha$ and

$$f(t) = \det([T]_\beta - tI) = \det(B_1 - tI) \det(B_3 - tI) = g(t)h(t).$$

29. We use the notation given in Hint of the previous exercise again. By Exercise 5.4.24 we may find a basis γ for W such that $[T_W]_\gamma$ is diagonal. For each eigenvalue λ, we may pick the corresponding eigenvectors in γ and extend it to be a basis for the corresponding eigenspace. By collecting all these basis, we may form a basis β for V who is union of these basis. So we know that $[T]_\beta$ is diagonal. Hence the matrix B_3 is also diagonal. Since we've proven that $B_3 = [\overline{T}]_\alpha$, we know that $\overline{T}$ is diagonalizable.

30. We also use the notation given in Hint of the previous exercise. By Exercise 5.4.24 we may find a basis

$$\gamma = \{v_1, v_2, \ldots, v_k\}$$

for W such that $[T_W]_\gamma$ is diagonal and find a basis

$$\alpha = \{v_{k+1} + W, v_{k+2} + W, \ldots, v_n + W\}$$

for V/W such that $[\overline{T}]_\alpha$ is diagonal. For each $v + W \in V/W$, we know that $v + W$ is an eigenvector in V/W. So we may assume that

$$T(v) + W = \overline{T}(v + W) = \lambda v + W$$

for some scalar λ. So this means that $T(v) - \lambda v$ is an element in W. Write

$$T(v) = \lambda v + a_1 v_1 + a_2 v_2 + \cdots + a_k v_k,$$

141

where v_i's are those elements in γ corresponding to eigenvalues λ_i's. Pick c_i to be $\frac{a_i}{\lambda - \lambda_i}$ and set

$$v' = v + c_1 v_1 + c_2 v_2 + \cdots + c_k v_k.$$

Those c_i's are well-defined because T_W and $\overline{T}$ has no common eigenvalue. Thus we have

$$T(v') = T(v) + c_1 T(v_1) + c_2 T(v_2) + \cdots + c_k T(v_k)$$

$$= \lambda v + (a_1 + c_1 \lambda_1)v_1 + (a_2 + c_2 \lambda_2)v_2 + \cdots + (a_k + c_k \lambda_k)v_k)$$

$$= \lambda v'$$

and $v + W = v' + W$. By doing this process, we may assume that v_j is an eigenvector in V for each $v_j + W \in \alpha$. Finally we pick

$$\beta = \{v_1, v_2, \ldots, v_n\}$$

and show that it's a basis for V. We have that γ is independent and

$$\delta = \{v_{k+1}, v_{k+2}, \ldots, v_n\}$$

is also independent since $\eta(\delta) = \alpha$ is independent, where η is the quotient mapping defined in Exercise 5.4.27. Finally we know that if

$$a_{k+1} v_{k+1} + a_{k+2} v_{k+2} + \cdots a_n v_n \in W,$$

then

$$\left(a_{k+1} v_{k+1} + a_{k+2} v_{k+2} + \cdots a_n v_n\right) + W = 0 + W$$

and so

$$a_{k+1} = a_{k+2} = \cdots = a_n = 0.$$

This means that $\operatorname{span}(\delta) \cap W = \{0\}$. So V is a directed sum of W and $\operatorname{span}(\delta)$ and β is a basis by Exercise 1.6.33. And thus we find a basis consisting of eigenvectors.

31. (a) Compute

$$T(e_1) = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, T^2(e_1) = \begin{pmatrix} 0 \\ 12 \\ 6 \end{pmatrix}$$

and

$$T^2(e_1) = -6e_1 + 6T(e_1).$$

This means that the characteristic polynomial of T_W is $t^2 - 6t + 6$.

(b) We know that
$$\dim(\mathbb{R}^3/W) = 3 - \dim(W) = 1$$

by Exercise 1.6.35. So every nonzero element in $\mathbb{R}^3/W$ is a basis. Since e_2 is not in W, we have $e_2 + W$ is a basis for $\mathbb{R}^3/W$. Now let $\beta = \{e_2 + W\}$. We may compute

$$\overline{T}(e_2 + W) = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} + W = -e_2 + W$$

and $[\overline{T}]_\beta = (1)$. So the characteristic polynomial of $\overline{T}$ is $-t - 1$.

(c) Use the result in Exercise 5.4.28, we know the characteristic polynomial of T is
$$-(t+1)(t^2 - 6t + t).$$

32. As Hint, we induct on the dimension n of the space. For $n = 1$, every matrix representation is an uppertriangular matrix. Suppose the hypothesis is true for $n = k - 1$. Now we consider the case $n = k$. Since the characteristic polynomial splits, T must has an eigenvector v corresponding to eigenvalue λ. Let $W = \text{span}(\{v\})$. By Exercise 1.6.35 we know that the dimension of V/W is equal to $k - 1$. So we apply the induction hypothesis to $\overline{T}$ defined in Exercise 5.4.27. This means we may find a basis

$$\alpha = \{u_2 + W, u_3 + W, \dots, u_k + W\}$$

for V/W such that $[\overline{T}]_\alpha$ is an upper triangular matrix. Following the argument in Exercise 5.4.30, we know that

$$\beta = \{v, u_2, u_3, \dots, u_k\}$$

is a basis for V. And we also know the matrix representation is

$$[T]_\beta = \begin{pmatrix} \lambda & * \\ O & [\overline{T}]_\alpha \end{pmatrix},$$

which is an upper triangular matrix since $[\overline{T}]_\alpha$ is upper triangular.

33. Let
$$w = w_1 + w_2 + \cdots + w_k$$

for some $w_i \in W_i$. Thus we have $T(w_i) \in W_i$ since W_i is T-invariant. So we also have

$$T(w) = T(w_1) + T(w_2) + \cdots + T(w_k) \in W_1 + W_2 + \cdots W_k.$$

34. In the next exercise we prove this exercise and the next exercise together.

35. Let v is an element in β_1. We know that $T(v)$ is a linear combination of β. Since $T(v)$ is in W_1, we know that the combination only use elements in β_1. The same argument could be applied to each element in β_i for each i. So we know the matrix representation of T corresponding to the chosen basis β is
$$B_1 \oplus B_2 \oplus \cdots \oplus B_k.$$

36. First observe that a one-dimensional T-invariant subspace is equivalent to a subspace spanned by an eigenvector. Second, V is the direct sum of some one-dimensional subspaces if and only if the set obtained by choosing one nonzero vector from each subspace forms a basis. Combining these two observation we get the result.

37. Use Theorem 5.25 and its notation. We have
$$\det(T) = \det(B) = \det(B_1)\det(B_2)\cdots\det(B_k)$$
$$= \det(T_{W_1})\det(T_{W_2})\cdots\det(T_{W_k}).$$

38. By Exercise 5.4.24, we already get the necessity. For the sufficiency, we may pick bases β_i for W_i such that $[T_{W_i}]$ is diagonal. Combining these bases to be a basis $\beta = \cup_{i=1}^{k}\beta_i$ for V. By using Theorem 5.25, we know that T is diagonalizable.

39. By Exercise 5.2.18, we already get the necessity. For the sufficiency, we use induction on the dimension n of the spaces. If $n = 1$, every operator on it is diagonalizable by the standard basis. By supposing the statement is true for $n \leq k-1$, consider the case $n = k$. First, if all the operator in $\mathcal{C}$ has only one eigenvalue, then we may pick any basis β and know that $[T]_\beta$ is diagonal for all $T \in \mathcal{C}$. Otherwise, there must be one operator T possessing two or more eigenvalues, $\lambda_1, \lambda_2, \ldots, \lambda_t$. Let W_i be the eigenspace of T corresponding to the eigenvalue λ_i. We know that V is the direct sum of all W_i's. By the same reason in the proof of Exercise 5.4.25, we know that W_i is a U-invariant subspace for all $U \in \mathcal{C}$. Thus we may apply the induction hypothesis on W_i and
$$\mathcal{C}_i := \{U_{W_i} : U \in \mathcal{C}\}.$$

Thus we get a basis β_i for W_i such that $[U_{W_i}]_{\beta_i}$ is diagonal. Let β be the union of each β_i. By applying Theorem 5.25, we get the desired result.

40. Observe that
$$A - tI = (B_1 - tI) \oplus (B_2 - tI) \oplus \cdots \oplus (B_k - tI).$$

So we have
$$\det(A - tI) = \prod_{i=1}^{k}\det(B_i - tI).$$

41. Let v_i be the i-th row vector of A. We have $\{v_1, v_2\}$ is linearly independent and $v_i = (i-1)(v_2 - v_1) + v_1$. So the rank of A is 2. This means that t^{n-2} is a factor of the characteristic polynomial of A. Finally, set

$$\beta = \{(1, 1, \ldots, 1), (1, 2, \ldots, n)\}$$

and check that $W = \text{span}(\beta)$ is a L_A-invariant subspace by computing

$$A \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = (\frac{n(n+1)}{2} - n^2) \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} + n^2 \begin{pmatrix} 1 \\ 2 \\ \vdots \\ n \end{pmatrix}$$

and

$$A \begin{pmatrix} 1 \\ 2 \\ \vdots \\ n \end{pmatrix} = (\frac{n(n+1)(2n+1)}{6} - \frac{n^2(n+1)}{2}) \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} + \frac{n^2(n+1)}{2} \begin{pmatrix} 1 \\ 2 \\ \vdots \\ n \end{pmatrix}.$$

So we know the characteristic polynomial is

$$(-1)^n t^{n-2} \det \begin{pmatrix} \frac{n(n+1)}{2} - n^2 - t & \frac{n(n+1)(2n+1)}{6} - \frac{n^2(n+1)}{2} \\ n^2 & \frac{n^2(n+1)}{2} - t \end{pmatrix}$$

$$= (-1)^n t^{n-2} \frac{12t^2 + (-6n^3 - 6n)t - n^5 + n^3}{12}.$$

But this is the formula for $n \geq 2$. It's natural that when $n = 1$ the characteristic polynomial is $1 - t$. Ur...I admit that I computed this strange answer by wxMaxima.

42. Observe that the nullity of A is $n - 1$ and $\det(A - nI) = 0$ because the sum of all column vectors of $A - nI$ is zero. So we know the characteristic polynomial of it is

$$(-1)^n t^{n-1}(t - n).$$

Chapter 6

Inner Product Spaces

6.1 Inner Products and Norms

1. (a) Yes. It's the definition.

   (b) Yes. See the paragraph at the beginning of this chapter.

   (c) No. It's conjugate linear in the second component. For example, in $\mathbb{C}$ with standard inner product function we have $\langle i, i \rangle = 1$ but not $i \langle i, 1 \rangle = -1$.

   (d) No. We may define the inner product function f on $\mathbb{R}$ to be $f(u, v) = 2uv$.

   (e) No. Theorem 6.2 does not assume that the dimension should be finite.

   (f) No. We may define the conjugate-transpose of any matrix A to be the conjugate of A^t.

   (g) No. Let $x = (1, 0)$, $y = (0, 1)$, and $z = (0, -1)$. Then we have $\langle x, y \rangle = \langle x, z \rangle = 0$.

   (h) Yes. This means $\|y\| = \langle y, y \rangle = 0$.

2. Compute
$$\langle x, y \rangle = 2 \cdot (2 + i) + (1 + i) \cdot 2 + i \cdot (1 - 2i) = 8 + 5i,$$

$$\|x\| = \langle x, x \rangle^{\frac{1}{2}} = (4 + 2 + 1)^{\frac{1}{2}} = \sqrt{7},$$

$$\|y\| = \langle y, y \rangle^{\frac{1}{2}} = (5 + 4 + 5)^{\frac{1}{2}} = \sqrt{14},$$

and
$$\|x + y\| = \|(4 - i, 3 + i, 1 + 3i)\| = (17 + 10 + 10)^{\frac{1}{2}} = \sqrt{37}.$$

We have the Cauchy-Schwarz inequality and the triangle inequality hold since
$$|8 + 5i| = \sqrt{89} \le \sqrt{7}\sqrt{14} = \sqrt{119}$$

and
$$\sqrt{7} + \sqrt{14} \geq \sqrt{37}.$$

3. As definition, compute
$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt = \int tde^t = te^t\big|_0^1 - \int e^t dt = e - (e - 1) = 1,$$

$$\|f\| = \int_0^1 f^2 = \frac{1}{3},$$

$$\|g\| = \int_0^1 g^2 = \frac{1}{2}(e^2 - 1),$$

and
$$\|f + g\| = \|t + e^t\| = (\|f\|^2 + 2\langle f, g \rangle + \|g\|^2)^{\frac{1}{2}}$$
$$= \frac{1}{9} + 2 + \frac{1}{4}(e^4 - 2e^2 + 1) = \frac{9e^4 - 18e^2 + 85}{36}.$$

And the two inequalities hold since
$$1 \leq \frac{1}{3} \cdot \frac{1}{2}(e^2 - 1)$$

and
$$\frac{9e^4 - 18e^2 + 85}{36} \geq \frac{1}{3} + \frac{1}{2}(e^2 - 1).$$

4. (a) We prove the formula
$$\langle A, B \rangle = \text{tr}(B^*A) = \sum_{j=1}^n (B^*A)_{jj} = \sum_{j=1}^n \sum_{i=1}^n B_{ji}^* A_{ij}$$

$$= \sum_{j=1}^n \sum_{i=1}^n A_{ij}\overline{B}_{ij} = \sum_{i=1}^n \sum_{j=1}^n A_{ij}\overline{B}_{ij} = \sum_{i,j} A_{ij}\overline{B}_{ij}.$$

So we may view the space $M_{n\times n}(\mathbb{F})$ to be F^{n^2} and the Frobenius inner product is corresponding to the standard inner product in F^{n^2}.

(b) Also use the formule to compute
$$\|A\| = (1 + 5 + 9 + 1)^{\frac{1}{2}} = 4,$$

$$\|B\| = (2 + 0 + 1 + 1)^{\frac{1}{2}} = 2,$$

and
$$\langle A, B \rangle = (1 - i) + 0 - 3i - 1 = -4i.$$

5. We prove the condition for an inner product space one by one.

-
$$\langle x + z, y \rangle = (x + z)Ay^* = xAY^* + zAy^* = \langle x, y \rangle + \langle z, y \rangle.$$

147

-
$$\langle cx, y \rangle = (cx)Ay^* = c(xAy^*) = c\langle x, y \rangle.$$

-
$$\overline{\langle x, y \rangle} = (xAy^*)^* = yA^*x^* = yAx^* = \langle y, x \rangle.$$

One of the equality use the fact $A = A^*$.

-
$$\langle x, x \rangle = (x_1, x_2)A(x_1, x_2)^* = \|x_1\|^2 + ix_1\overline{x_2} - ix_2\overline{x_1} + 2\|x_2\|^2$$
$$= \|x_1\|^2 + 2\mathrm{Re}(ix_1\overline{x_2}) + 2\|x_2\|^2 \rangle 0$$

if x_1 or x_2 is not 0. Here the function $\mathrm{Re}(z)$ means the real part of a complex number z.

So it's an inner product function. Also compute that

$$\langle x, y \rangle = 1(1-i)(2-i) + i(1-i)(3+2i) + (-i)(2+3i)(2-i) + 2(2+3i)(3+2i)$$
$$= (1-i)(2i) + (2+3i)(5+2i) = 6 + 21i.$$

6. (a)
$$\langle x, y+z \rangle = \overline{\langle y+z, x \rangle} = \overline{\langle y, x \rangle} + \overline{\langle z, x \rangle} = \langle x, y \rangle + \langle x, z \rangle.$$

   (b)
$$\langle x, cy \rangle = \overline{\langle cy, x \rangle} = \overline{c\langle y, x \rangle} = \overline{c}\langle x, y \rangle.$$

   (c)
$$\langle x, 0 \rangle = \overline{0}\langle x, 0 \rangle = 0$$

   and
$$\langle 0, x \rangle = \overline{\langle x, 0 \rangle} = 0.$$

   (d) If $x = 0$, then $\langle 0, 0 \rangle = 0$ by previous rule. If $x \neq 0$, then $\langle x, x \rangle > 0$.

   (e) If $\langle x, y \rangle = \langle x, z \rangle$ for all $x \in V$, we have $\langle x, y - z \rangle = 0$ for all $x \in V$. So we have $\langle y - z, y - z \rangle = 0$ and hence $y - z = 0$.

7. (a)
$$\|cx\| = \langle cx, cx \rangle^{\frac{1}{2}} = (c\overline{c}\langle x, x \rangle)^{\frac{1}{2}} = |c|\|x\|.$$

   (b) This is the result of $\|x\| = \langle x, x \rangle^{\frac{1}{2}}$ and Theorem 6.1(d).

8. (a) The inner product of a nonzero vector $(1, 1)$ and itself is $1 - 1 = 0$.

   (b) Let $A = B = I_2$. We have

$$\langle 2A, B \rangle = 3 \neq 2\langle A, B \rangle = 4.$$

   (c) The inner product of a nonzero function $f(x) = 1$ and itself is

$$\int_0^1 0 \cdot 1 dx = 0.$$

148

9. (a) Represent x to be a linear combination of β as

$$x = \sum_{i=1}^{k} a_i z_i,$$

where z_i's are elements in β. Then we have

$$\langle x, x \rangle = \langle x, \sum_{i=1}^{k} a_i z_i \rangle$$

$$= \sum_{i=1}^{k} \overline{a_i} \langle x, z_i \rangle = 0$$

and so $x = 0$.

(b) This means that $\langle x - y, z \rangle = 0$ for all $z \in \beta$. So we have $x - y = 0$ and $x = y$.

10. Two vectors are orthogonal means that the inner product of them is 0. So we have

$$\|x + y\|^2 = \langle x + y, x + y \rangle$$

$$= \|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2 = \|x\|^2 + + \|y\|^2.$$

To deduce the Pythagorean Theorem in R^2, we may begin by a right triangle ABC with the right angle B. Assume that $x = AB$ and $y = BC$. Thus we know the length of two leg is $\|x\|$ and $\|y\|$. Finally we know $AC = x + y$ and so the length of the hypotenuse is $\|x + y\|$. Apply the proven equality and get the desired result.

11. Compute that

$$\|x + y\|^2 + \|x - y\|^2$$

$$= (\|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2) + (\|x\|^2 - \langle x, y \rangle - \langle y, x \rangle + \|y\|^2)$$

$$= 2\|x\|^2 + 2\|y\|^2.$$

This means that the sum of square of the four edges of a prallelogram is the sum of square of the two diagonals.

12. Compute that

$$\| \sum_{i=1}^{k} a_i v_i \|^2 = \langle \sum_{i=1}^{k} a_i v_i \rangle$$

$$= \sum_{i=1}^{k} \|a_i v_i\|^2 + \sum_{i,j} \langle a_i v_i, a_j v_j \rangle$$

$$= \sum_{i=1}^{k} |a_i|^2 \|v_i\|.$$

13. Check those condition.

-
$$\langle x + z, y \rangle = \langle x + z, y \rangle_1 + \langle x + z, y \rangle_2$$
$$= \langle x, y \rangle_1 + \langle z, y \rangle_1 + \langle x, y \rangle_2 + \langle z, y \rangle_2$$
$$= \langle x, y \rangle + \langle z, y \rangle.$$

-
$$\langle cx, y \rangle = \langle cx, y \rangle_1 + \langle cx, y \rangle_2$$
$$= c\langle x, y \rangle_1 + c\langle x, y \rangle_2 = c\langle x, y \rangle.$$

-
$$\overline{\langle x, y \rangle} = \overline{\langle x, y \rangle_1} + \overline{\langle x, y \rangle_2}$$
$$= \langle y, x \rangle_1 + \langle y, x \rangle_2 = \langle y, x \rangle.$$

-
$$\langle x, x \rangle = \langle x, x \rangle_1 + \langle x, x \rangle_2 > 0$$

if $x \neq 0$.

14. Check that

$$(A + cB)^*_{ij} = \overline{(A + cB)_{ji}} = \overline{A_{ji}} + \overline{c}\,\overline{B_{ji}} = (A^* + \overline{c}B^*)_{ij}.$$

15. (a) If one of x or y is zero, then the equality holds naturally and we have $y = 0x$ or $x = 0y$. So we may assume that y is not zero. Now if $x = cy$, we have

$$|\langle x, y \rangle| = |\langle cy, y \rangle| = |c| \|y\|^2$$

and

$$\|x\| \cdot \|y\| = \|cy\| \cdot \|y\| = |c| \|y\|^2.$$

For the necessity, we just observe the proof of Theorem 6.2(c). If the equality holds, then we have

$$\|x - cy\| = 0,$$

where

$$c = \frac{\langle x, y \rangle}{\langle y, y \rangle}.$$

And so $x = cy$.

(b) Also observe the proof of Theorem 6.2(d). The equality holds only when

$$\mathrm{Re}\langle x, y \rangle = |\langle x, y \rangle| = \|x\| \cdot \|y\|.$$

The case $y = 0$ is easy. Assuming $y \neq 0$, we have $x = cy$ for some scalar $c \in F$. And thus we have

$$\mathrm{Re}(c)\|y\|^2 = \mathrm{Re}\langle cy, y \rangle = |\langle cy, y \rangle| = |c| \cdot \|y\|^2$$

and so
$$\mathrm{Re}(c) = |c|.$$

This means c is a nonnegative real number. Conversely, if $x = cy$ for some nonnegative real number, we may check

$$\|x + y\| = |c + 1|\|y\| = (c + 1)\|y\|$$

and

$$\|x\| + \|y\| = |c|\|y\| + \|y\| = (c + 1)\|y\|.$$

Finally, we may generalize it to the case of n vectors. That is,

$$\|x_1 + x_2 + \cdots + x_n\| = \|x_1\| + \|x_2\| + \cdots + \|x_n\|$$

if and only if we may pick one vector from them and all other vectors are some multiple of that vector using nonnegative real number.

16. (a) Check the condition one by one.

- $$\langle f + h, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} (f(t) + h(t))\overline{g(t)}dt$$
$$= \frac{1}{2\pi} \int_0^{2\pi} f\overline{g}dt + \frac{1}{2\pi} \int_0^{2\pi} (h\overline{g}dt = \langle f, g \rangle + \langle h, g \rangle.$$

- $$\langle cf, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} cf\overline{g}dt$$
$$= c(\frac{1}{2\pi} \int_0^{2\pi} f\overline{g}dt) = c\langle f, g \rangle.$$

- $$\overline{\langle f, g \rangle} = \overline{\frac{1}{2\pi} \int_0^{2\pi} f\overline{g}dt} = \frac{1}{2\pi} \int_0^{2\pi} \overline{f\overline{g}}dt$$
$$= \frac{1}{2\pi} \int_0^{2\pi} g\overline{f}dt = \langle g, f \rangle.$$

- $$\langle f, f \rangle = \frac{1}{2\pi} \int_0^{2\pi} \|f\|^2 dt > 0$$

  if f is not zero. Ur... I think this is an exercise for the Adavanced Calculus course.

  (b) No. Let
  $$f(t) = \begin{cases} 0 & \text{if} \quad x \le \frac{1}{2}; \\ x - \frac{1}{2} & \text{if} \quad x > \frac{1}{2}. \end{cases}$$

  Then we have that $\langle f, f \rangle = 0$ but $f \ne 0$.

17. If $T(x) = 0$ Then we have
$$\|x\| = \|0\| = 0.$$

This means $x = 0$ and so T is injective.

18. If $\langle \cdot, \cdot \rangle'$ is an inner product on V, then we have that $T(x) = 0$ implies

$$\langle x, x \rangle' = \langle T(x), T(x) \rangle = 0$$

and $x = 0$. So T is injective. Conversely, if T is injective, we may check those condition for inner product one by one.

*
$$\langle x + z, y \rangle' = \langle T(x + z), T(y) \rangle = \langle T(x) + T(z), T(y) \rangle$$
$$= \langle T(x), T(y) \rangle + \langle T(z), T(y) \rangle = \langle x, y \rangle' + \langle z, y \rangle'.$$

*
$$\langle cx, y \rangle' = \langle T(cx), T(y) \rangle = \langle cT(x), T(y) \rangle$$
$$= c \langle T(x), T(y) \rangle = \langle x, y \rangle'.$$

*
$$\overline{\langle x, y \rangle'} = \overline{\langle T(x), T(y) \rangle}$$
$$= \langle T(y), T(x) \rangle = \langle y, x \rangle'.$$

*
$$\langle x, x \rangle' = \langle T(x), T(x) \rangle > 0$$

  if $T(x) \neq 0$. And the condition $T(x) \neq 0$ is true when $x \neq 0$ since T is injective.

19. (a) Just compute

$$\|x + y\|^2 = \|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2$$

$$= \|x\|^2 + \langle x, y \rangle + \overline{\langle y, x \rangle} + \|y\|^2 = \|x\|^2 + \mathrm{Re}(\langle x, y \rangle) + \|y\|^2$$

and
$$\|x - y\|^2 = \|x\|^2 - \langle x, y \rangle - \langle y, x \rangle + \|y\|^2$$
$$= \|x\|^2 - \langle x, y \rangle - \overline{\langle y, x \rangle} + \|y\|^2 = \|x\|^2 - \mathrm{Re}(\langle x, y \rangle) + \|y\|^2.$$

   (b) We have
$$\|x - y\| + \|y\| \geq \|x\|$$

and
$$\|x - y\| + \|x\| = \|y - x\| + \|x\| \geq \|y\|.$$

Combining these two we get the desired inequality.

20. (a) By Exercise 6.1.19(a), we have the right hand side would be

$$\frac{1}{4}(4\mathrm{Re}\langle x, y \rangle) = \langle x, y \rangle.$$

The last equality is due to the fact $\mathbb{F} = \mathbb{R}$.

152

(b) Also use the Exercise 6.1.19(a). First observe that

$$\|x + i^k y\|^2 = \|x\|^2 + 2\mathrm{Re}\langle x, i^k y\rangle + \|i^k y\|^2$$

$$= \|x\|^2 + 2\mathrm{Re}[\overline{i^k}\langle x, y\rangle] + \|y\|^2.$$

Assuming $\langle x, y\rangle$ to be $a + bi$, we have the right hand side would be

$$\frac{1}{4}[(\sum_{k=1}^{4} i^k)\|x\|^2 + 2\sum_{k=1}^{4} i^k \mathrm{Re}[\overline{i^k}\langle x, y\rangle] + (\sum_{k=1}^{4} i^k)\|y\|^2]$$

$$= \frac{1}{2}[bi + (-a)(-1) + (-b)(-i) + a] = a + bi = \langle x, y\rangle.$$

21. (a) Observe that

$$A_1^* = (\frac{1}{2}(A + A^*))^* = \frac{1}{2}(A^* + A) = A_1^*,$$

$$A_2^* = (\frac{1}{2i}(A - A^*))^* = (-\frac{1}{2i}(A^* - A) = A_2^*.$$

and

$$A_1 + iA_2 = \frac{1}{2}(A + A^*) + \frac{1}{2}(A - A^*) = \frac{1}{2}(A + A) = A.$$

I don't think it's reasonable because A_1 does not consists of the real part of all entries of A and A_2 does not consists of the imaginary part of all entries of A. But what's the answer do you want? Would it be reasonable to ask such a strange question?

(b) If we have $A = B_1 + iB_2$ with $B_1^* = B_1$ and $B_2^* = B_2$, then we have $A^* = B_1^* - iB_2^*$. Thus we have

$$B_1 = \frac{1}{2}(A + A^*)$$

and

$$B_2 = \frac{1}{2i}(A - A^*).$$

22. (a) As definition, we may find $v_1, v_2, \ldots, v_n \in \beta$ for $x, y, z \in V$ such that

$$x = \sum_{i=1}^{n} a_i v_i, y = \sum_{i=1}^{n} b_i v_i, z = \sum_{i=1}^{n} d_i v_i.$$

And check those condition one by one.

●

$$\langle x + z, y\rangle = \langle \sum_{i=1}^{n} (a_i + d_i)v_i, \sum_{i=1}^{n} b_i v_i \rangle$$

$$= \sum_{i=1}^{n} (a_i + d_i)\overline{b_i} = \sum_{i=1}^{n} a_i \overline{b_i} + \sum_{i=1}^{n} d_i \overline{b_i}$$

$$= \langle x, y\rangle + \langle z, y\rangle.$$

-
$$\langle cx, y \rangle = \langle \sum_{i=1}^{n} (ca_i)v_i, \sum_{i=1}^{n} b_i v_i \rangle$$

$$= \sum_{i=1}^{n} (ca_i)\overline{b_i} = c \sum_{i=1}^{n} a_i \overline{b_i} = c\langle x, y \rangle.$$

-
$$\overline{\langle x, y \rangle} = \overline{\sum_{i=1}^{n} a_i \overline{b_i}}$$

$$= \sum_{i=1}^{n} \overline{a_i \overline{b_i}} = \sum_{i=1}^{n} b_i \overline{a_i} = \langle y, x \rangle.$$

-
$$\langle x, x \rangle = \sum_{i=1}^{n} |a_i|^2 > 0$$

    if all a_i's is not zero. That is, x is not zero.

(b) If the described condition holds, for each vector

$$x = \sum_{i=1}^{n} a_i v_i$$

we have actually a_i is the i-th entry of x. So the function is actually the atandard inner product. Note that this exercise give us an idea that different basis will give a different inner product.

23. (a) We have the fact that with the standard inner product $\langle \cdot, \cdot \rangle$ we have $\langle x, y \rangle = y^* x$. So we have

$$\langle x, Ay \rangle = (Ay)^* x = y^* A^* x = \langle A^* x, y \rangle.$$

(b) First we have that

$$\langle A^* x, y \rangle = \langle x, Ay \rangle = \langle Bx, y \rangle$$

for all $x, y \in V$. By Theorem 6.1(e) we have $A^* x = Bx$ for all x. But this means that these two matrix is the same.

(c) Let $\beta = \{v_1, v_2, \ldots, v_n\}$. So the column vectors of Q are those v_i's. Finally observe that $(Q^* Q)_{ij}$ is

$$v_i^* v_j = \langle v_i, v_j \rangle = \begin{cases} 1 & \text{if} \quad i = j; \\ 0 & \text{if} \quad i \neq j. \end{cases}$$

So we have $Q^* Q = I$ and $Q^* = Q^{-1}$.

(d) Let α be the standard basis for $\mathbb{F}^n$. Thus we have $[T]_\alpha = A$ and $[U]_\alpha = A^*$. Also we have that actually $[I]_\alpha^\beta$ is the matrix Q defined in the previous exercise. So we know that

$$[U]_\beta = [I]_\alpha^\beta [U]_\alpha [I]_\beta^\alpha = QAQ^{-1} = QAQ^*$$

$$= (QA^* Q^*)^* = [I]_\alpha^\beta [T]_\alpha [I]_\beta^\alpha = [T]_\beta.$$

154

24. Check the three conditions one by one.

(a) •
$$\|A\| = \max_{i,j} |A_{ij}| \geq 0,$$

and the value equals to zero if and only if all entries of A are zero.

•
$$\|aA\| = \max_{i,j} |(aA)_{ij}| = \max_{i,j} |a||A_{ij}|$$
$$= |a| \max_{i,j} |A_{ij}| = |a|\|A\|.$$

•
$$\|A + B\| = \max_{i,j} |(A+B)_{ij}| = \max_{i,j} |A_{ij} + B_{ij}|$$
$$\leq \max_{i,j} |A_{ij}| + \max_{i,j} |B_{ij}| = \|A\| + \|B\|.$$

(b) •
$$\|f\| = \max_{t\in[0,1]} |f(t)| \geq 0,$$

and the value equals to zero if and only if all value of f in $[0,1]$ is zero.

•
$$\|af\| = \max_{t\in[0,1]} |(af)(t)| = \max_{t\in[0,1]} |a||f(t)|$$
$$= |a| \max_{t\in[0,1]} |f(t)| = |a|\|f\|.$$

•
$$\|f + g\| = \max_{t\in[0,1]} |(f+g)(t)| = \max_{t\in[0,1]} |f(t) + g(t)|$$
$$\leq \max_{t\in[0,1]} |f(t)| + \max_{t\in[0,1]} |g(t)| = \|f\| + \|g\|.$$

(c) •
$$\|f\| = \int_0^1 |f(t)|dt \geq 0,$$

and the value equals to zero if and only if $f = 0$. This fact depend on the continuity and it would be an exercise in the Advanced Calculus coures.

•
$$\|af\| = \int_0^1 |af(t)|dt = \int_0^1 |a||f(t)|dt$$
$$= |a| \int_0^1 |f(t)| = |a|\|f\|.$$

-
$$\|f + g\| = \int_0^1 |f(t) + g(t)|dt \le \int_0^1 |f(t)| + |g(t)|dt$$
$$= \int_0^1 |f(t)|dt + \int_0^1 |g(t)|dt = \|f\| + \|g\|.$$

(d)
-
$$\|(a,b)\| = \max\{|a|, |b|\} \ge 0,$$
and the value equals to zero if and only if both a and b are zero.

-
$$\|c(a,b)\| = \max\{|ca|, |cb|\} = \max\{|c||a|, |c||b|\}$$
$$= |c| \max\{|a|, |b|\} = |c| \|(a,b)\|.$$

-
$$\|(a,b) + (c,d)\| = \max\{|a + c|, |b + d|\} \le \max\{|a| + |c|, |b| + |d|\}$$
$$\le \max\{|a|, |b|\} + \max\{|c|, |d|\} = \|(a,b)\| + \|(c,d)\|.$$

25. By Exercise 6.1.20 we know that if there is an inner product such that $\|x\|^2 = \langle x, x\rangle$ for all $x \in \mathbb{R}^2$, then we have

$$\langle x, y\rangle = \frac{1}{4}\|x + y\|^2 - \frac{1}{4}\|x - y\|.$$

Let $x = (2,0)$ and $y = (1,3)$. Thus we have

$$\langle x, y\rangle = \frac{1}{4}\|(3,3)\|^2 - \frac{1}{4}\|(1,-3)\|^2 = 0$$

and
$$\langle 2x, y\rangle = \frac{1}{4}\|(5,3)\|^2 - \frac{1}{4}\|(3,-3)\|^2 = \frac{1}{4}(25 - 9) = 4.$$

This means this function is not linear in the first component.

26. (a)
$$d(x,y) = \|x - y\| \ge 0.$$

(b)
$$d(x,y) = \|x - y\| = \|y - x\| = d(y,x).$$

(c)
$$d(x,y) = \|x - y\| = \|(x - z) + (z - y)\| \le \|x - z\| + \|z - y\| = d(x,y) + d(z,y).$$

(d)
$$d(x,x) = \|x - x\| = \|0\| = 0.$$

(e)
$$d(x, y) = \|x - y\| > 0$$
if $x - y$ is not zero. That is, the distance is not zero if $x \neq y$.

27. The third and the fourth condition of inner product naturally hold since
$$\langle x, y \rangle = \frac{1}{4}[\|x + y\|^2 - \|x - y\|^2] = \langle y, x \rangle$$
and
$$\langle x, x \rangle = \frac{1}{4}[\|2x\|^2 - \|0\|^2] = \|x\|^2 > 0$$
if $x \neq 0$. Now we prove the first two condition as Hints.

(a) Consider that
$$\|x + 2y\|^2 + \|x\|^2 = 2\|x + y\|^2 + 2\|y\|^2$$
and
$$\|x - 2y\|^2 + \|x\|^2 = 2\|x - y\|^2 + 2\|y\|^2$$
by the parallelogram law. By Substracting these two equalities we have
$$\|x + 2y\|^2 - \|x - 2y\|^2 = 2\|x + y\|^2 - 2\|x - y\|^2.$$
And so we have
$$\langle x, 2y \rangle = \frac{1}{4}[\|x + 2y\|^2 - \|x - 2y\|^2]$$
$$= \frac{1}{4}[2\|x + y\|^2 - 2\|x - y\|^2] = 2\langle x, y \rangle.$$

(b) By the previous argument, we direct prove that
$$\langle x + u, 2y \rangle = 2\langle x, y \rangle + 2\langle u, y \rangle.$$
Similarly begin by
$$\|x + u + 2y\|^2 + \|x - u\|^2 = 2\|x + y\|^2 + 2\|u + y\|^2$$
and
$$\|x + u - 2y\|^2 + \|x - u\|^2 = 2\|x - y\|^2 + 2\|u - y\|^2$$
by the parallelogroom law. By substracting these two equalities we have
$$\|x + u + 2y\|^2 - \|x + u - 2y\|^2 = 2\|x + y\|^2 - 2\|x - y\|^2 + 2\|u + y\|^2 - 2\|u - y\|^2.$$
And so we have
$$\langle x + u, 2y \rangle = \frac{1}{4}[\|x + u + 2y\|^2 - \|x + u - 2y\|^2]$$
$$= \frac{1}{4}[2\|x + y\|^2 - 2\|x - y\|^2 + 2\|u + y\|^2 - 2\|u - y\|^2]$$
$$= 2\langle x, y \rangle + 2\langle u, y \rangle.$$

157

(c) Since n is a positive integer, we have

$$\langle nx, y \rangle = \langle (n-1)x, y \rangle + \langle x, y \rangle$$

$$= \langle (n-2)x, y \rangle + 2\langle x, y \rangle = \cdots = n\langle x, y \rangle$$

by the previous argument inductively.

(d) Since m is a positive integer, we have

$$\langle x, y \rangle = \langle m(\frac{1}{m}x), y \rangle = m\langle \frac{1}{m}x, y \rangle$$

by the previous argument.

(e) Let $r = \frac{p}{q}$ for some positive integers p and q if r is positive. In this case we have

$$\langle rx, y \rangle = \langle p(\frac{1}{q}x), y \rangle = p\langle \frac{1}{q}x, y \rangle$$

$$= \frac{p}{q}\langle x, y \rangle = r\langle x, y \rangle.$$

If r is zero, then it's natural that

$$\langle 0, y \rangle = \frac{1}{4}[\|y\|^2 - \|y\|^2] = 0 = 0\langle x, y \rangle.$$

Now if r is negative, then we also have

$$\langle rx, y \rangle = \langle (-r)(-x), y \rangle = -r\langle -x, y \rangle.$$

But we also have

$$\langle -x, y \rangle = -\langle x, y \rangle$$

since

$$\langle -x, y \rangle + \langle x, y \rangle$$

$$= \frac{1}{4}[\| -x + y\|^2 - \| -x - y\|^2 + \|x + y\|^2 - \|x - y\|^2] = 0.$$

So we know that even when r is negative we have

$$\langle rx, y \rangle = -r\langle -x, y \rangle = r\langle x, y \rangle.$$

(f) Now we have the distribution law. Also, the position of the two component can be interchanged without change the value of the defined function. Finally we observe that

$$\langle x, x \rangle = \frac{1}{4}[\|x + x\|^2 - \|0\|^2] = \|x\|^2$$

for all $x \in V$. So now we have

$$\|x\|^2 + 2\langle x, y \rangle + \|y\|^2 = \|x + y\|^2$$

$$\leq (\|x\| + \|y\|)^2 = \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2$$

by the triangle inequality. Similarly we also have

$$\|x\|^2 - 2\langle x, y \rangle + \|y\|^2 = \|x - y\|^2$$

$$\leq (\|x\| + \| - y\|)^2 = \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2.$$

Thus we get the desired result

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|.$$

(g) Since

$$(c - r)\langle x, y \rangle = c\langle x, y \rangle - r\langle x, y \rangle$$

and

$$\langle (c - r)x, y \rangle = \langle cx - rx, y \rangle = \langle cx, y \rangle - r\langle x, y \rangle,$$

the first equality holds. And by the previous argument we have

$$-|c - r|\|x\|\|y\| \leq (c - r)\langle x, y \rangle, \langle (c - r)x, y \rangle \leq |c - r|\|x\|\|y\|$$

and so we get the final inequality.

(h) For every real number c, we could find a rational number such that $|c - r|$ is small enough[1]. So by the previous argument, we have

$$\langle cx, y \rangle = c\langle x, y \rangle$$

for all real number c.

28. Check the conditions one by one.

- $$[x + z, y] = \mathrm{Re}\langle x + z, y \rangle = \mathrm{Re}(\langle x, y \rangle + \langle z, y \rangle)$$
  $$= \mathrm{Re}\langle x, y \rangle + \mathrm{Re}\langle z, y \rangle = [x, y] + [z, y].$$

- $$[cx, y] = \mathrm{Re}\langle cx, y \rangle$$
  $$= c\mathrm{Re}\langle x, y \rangle = c[x, y],$$

  where c is a real number.

- $$[x, y] = \mathrm{Re}\langle x, y \rangle = \mathrm{Re}\overline{\langle x, y \rangle}$$
  $$= \mathrm{Re}\langle y, x \rangle = [y, x].$$

- $$[x, x] = \mathrm{Re}\langle x, x \rangle = \langle x, x \rangle > 0$$

  if $x \neq 0$.

---

[1] This is also an exercise for the Adavanced Calculus course.

Finally, we have $[x, ix] = 0$ since

$$\langle x, ix \rangle = -i \langle x, x \rangle$$

is a pure imaginary number.

29. Observe that

$$0 = [x + iy, i(x + iy)] = [x + iy, ix - y] = [ix, iy] - [x, y].$$

So we have an instant property

$$[x, y] = [ix, iy].$$

Now check the conditions one by one.

- $$\langle x + z, y \rangle = [x + z, y] + i[x + z, iy]$$
$$= [x, y] + [z, y] + i[x, iy] + i[z, iy] = \langle x, y \rangle + \langle z, y \rangle.$$

- $$\langle (a + bi)x, y \rangle = [(a + bi)x, y] + i[(a + bi)x, iy]$$
$$= [ax, y] + [bix, y] + i[ax, iy] + i[bix, iy]$$
$$= a([x, y] + i[x, iy]) + bi([ix, iy] - i[ix, y])$$
$$= a \langle x, y \rangle + bi([x, y] + i[x, iy]) = (a + bi)\langle x, y \rangle.$$

Here we use the proven property.

- $$\overline{\langle x, y \rangle} = [x, y] - i[x, iy]$$
$$= [y, x] + i[y, ix] = \langle x, y \rangle.$$

Here we use the proven property again.

- $$\langle x, x \rangle = [x, x] + i[x, ix] = [x, x] > 0$$

if x is not zero.

30. First we may observe that the condition for norm on real vector space is loosen than that on complex vector space. So naturally the function $\| \cdot \|$ is still a norm when we regard V as a vector space over $\mathbb{R}$. By Exercise 6.1.27, we've already defined a real inner product $[\cdot, \cdot]$ on it since the parallelogram law also holds on it. And we also have

$$[x, ix] = \frac{1}{4}[\|x + ix\|^2 - \|x - ix\|^2] =$$

$$= \frac{1}{4}[\|x + ix\|^2 - \|(-i)(x + ix)\|^2] = \frac{1}{4}[\|x + ix\|^2 - |-i|\|(x + ix)\|^2] = 0.$$

So by Exercise 6.1.29 we get the desired conclusion.

6.2 The Gram-Schmidt Orthogonalization Process and Orthogonal Complements

1. (a) No. It should be at least an independent set.

   (b) Yes. See Theorem 6.5.

   (c) Yes. Let W be a subspace. If x and y are elements in $W^\perp$ and c is a scalar, we have
   $$\langle x + y, w \rangle = \langle x, w \rangle + \langle y, w \rangle = 0$$
   and
   $$\langle cx, w \rangle = c\langle x, w \rangle = 0$$
   for all w in W. Furthermore, we also have $\langle 0, w \rangle = 0$ for all $w \in W$.

   (d) No. The basis should be orthonormal.

   (e) Yes. See the definition of it.

   (f) No. The set $\{0\}$ is orthogonal but not independent.

   (g) Yes. See the Corollary 2 after Theorem 6.3.

2. The answers here might be different due to the different order of vectors chosen to be orthogonalized.

   (a) Let
   $$S = \{w_1 = (1, 0, 1), w_2 = (0, 1, 1), w_3 = (1, 3, 3)\}.$$

   Pick $v_1 = w_1$. Then construct
   $$v_2 = w_2 - \frac{\langle w_2, v_1 \rangle}{\|v_1\|^2} v_1$$
   $$= w_2 - \frac{1}{2}v_1 = (-\frac{1}{2}, 1, \frac{1}{2}).$$

   And then construct
   $$v_3 = w_3 - \frac{\langle w_3, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle w_3, v_2 \rangle}{\|v_2\|^2} v_2$$
   $$= w_3 - \frac{4}{2}v_1 - \frac{4}{\frac{3}{2}}v_2 = (\frac{1}{3}, \frac{1}{3}, -\frac{1}{3}).$$

   As the demand in the exercise, we normalize v_1, v_2, and v_3 to be
   $$u_1 = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}),$$
   $$u_2 = (-\frac{1}{\sqrt{6}}, \frac{4}{\sqrt{6}}, \frac{1}{\sqrt{6}}),$$
   $$u_3 = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}).$$

Let $\beta = \{u_1, u_2, u_3\}$. Now we have two ways to compute the Fourier coefficients of x relative to β. One is to solve the system of equations

$$a_1 u_1 + a_2 u_2 + a_3 u_3 = x$$

and get

$$x = \frac{3}{\sqrt{2}} u_1 + \frac{3}{\sqrt{6}} u_2 + 0 u_3.$$

The other is to calculate the i-th Fourier coefficient

$$a_i = \langle x, u_i \rangle$$

directly by Theorem 6.5. And the two consequences meet.

(b) Ur...don't follow the original order. Pick $w_1 = (0, 0, 1)$, $w_2 = (0, 1, 1)$, and $w_3 = (1, 1, 1)$ and get the answer

$$\beta = \{(0, 0, 1), (0, 1, 0), (1, 0, 0)\}$$

instantly. And easily we also know that the Fourier coefficients of x relative to β are $1, 0, 1$.

(c) The basis is

$$\beta = \{1, \sqrt{3}(2x - 1), \sqrt{5}(6x^2 - 6x + 1)\}.$$

And the Fourier coefficients are $\frac{3}{2}, \frac{\sqrt{3}}{6}, 0$.

(d) The basis is

$$\beta = \{\frac{1}{\sqrt{2}}(1, i, 0), \frac{1}{2\sqrt{17}}(1 + i, 1 - i, 4i)\}.$$

And the Fourier coefficients are $\frac{7+i}{\sqrt{2}}, \sqrt{17}i$.

(e) The basis is

$$\beta = \{(\frac{2}{5}, -\frac{1}{5}, -\frac{2}{5}, \frac{4}{5}), (-\frac{4}{\sqrt{30}}, \frac{2}{\sqrt{30}}, -\frac{3}{\sqrt{30}}, \frac{1}{\sqrt{30}}),$$

$$(-\frac{3}{\sqrt{155}}, \frac{4}{\sqrt{155}}, \frac{9}{\sqrt{155}}, \frac{7}{\sqrt{155}})\}.$$

And the Fourier coefficients are $10, 3\sqrt{30}, \sqrt{150}$.

(f) The basis is

$$\beta = \{(\frac{1}{\sqrt{15}}, -\frac{2}{\sqrt{15}}, -\frac{1}{\sqrt{15}}, \frac{3}{\sqrt{15}}), (\frac{2}{\sqrt{10}}, \frac{2}{\sqrt{10}}, \frac{1}{\sqrt{10}}, \frac{1}{\sqrt{10}}),$$

$$(-\frac{4}{\sqrt{30}}, \frac{2}{\sqrt{30}}, \frac{1}{\sqrt{30}}, \frac{3}{\sqrt{30}})\}.$$

And the Fourier coefficients are $-\frac{3}{\sqrt{15}}, \frac{4}{\sqrt{10}}, \frac{12}{\sqrt{30}}$.

(g) The basis is

$$\beta = \left\{ \begin{pmatrix} \frac{1}{2} & \frac{5}{6} \\ -\frac{1}{6} & \frac{1}{6} \end{pmatrix}, \begin{pmatrix} -\frac{\sqrt{2}}{3} & \frac{\sqrt{2}}{3} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} \\ \frac{\sqrt{2}}{3} & -\frac{\sqrt{2}}{3} \end{pmatrix} \right\}.$$

And the Fourier coefficients are $24, 6\sqrt{2}, -9\sqrt{2}$.

(h) The basis is

$$\beta = \left\{ \begin{pmatrix} \frac{2}{\sqrt{13}} & \frac{2}{\sqrt{13}} \\ \frac{2}{\sqrt{13}} & \frac{1}{\sqrt{13}} \end{pmatrix}, \begin{pmatrix} \frac{5}{7} & -\frac{2}{7} \\ -\frac{4}{7} & \frac{2}{7} \end{pmatrix}, \begin{pmatrix} \frac{8}{\sqrt{373}} & -\frac{8}{\sqrt{373}} \\ \frac{7}{\sqrt{373}} & -\frac{14}{\sqrt{373}} \end{pmatrix} \right\}.$$

And the Fourier coefficients are $5\sqrt{13}, -14, \sqrt{373}$.

(i) The basis is

$$\beta = \left\{ \frac{\sqrt{2}\sin(t)}{\sqrt{\pi}}, \frac{\sqrt{2}\cos(t)}{\sqrt{\pi}}, \frac{\pi - 4\sin(t)}{\sqrt{\pi^3 - 8\pi}}, \frac{8\cos(t) + 2\pi t - \pi^2}{\sqrt{\frac{\pi^5}{3} - 32\pi}} \right\}.$$

And the Fourier coefficients are $\frac{\sqrt{2}(2\pi+2)}{\sqrt{\pi}}, -\frac{4\sqrt{2}}{\sqrt{\pi}}, \frac{\pi^3+\pi^2-8\pi-8}{\sqrt{\pi^3-8\pi}}, \frac{\frac{\pi^4-48}{3}-16}{\sqrt{\frac{\pi^5}{3}-32\pi}}$.

(j) The basis is

$$\left\{ \left(\frac{1}{2^{\frac{3}{2}}}, \frac{i}{2^{\frac{3}{2}}}, \frac{2-i}{2^{\frac{3}{2}}}, -\frac{1}{2^{\frac{3}{2}}} \right), \left(\frac{3i+1}{2\sqrt{5}}, \frac{i}{\sqrt{5}}, -\frac{1}{2\sqrt{5}}, \frac{2i+1}{2\sqrt{5}} \right), \right.$$

$$\left. \left(\frac{i-7}{2\sqrt{35}}, \frac{i+3}{\sqrt{35}}, \frac{5}{2\sqrt{35}}, \frac{5}{2\sqrt{35}} \right) \right\}.$$

And the Fourier coefficients are $6\sqrt{2}, 4\sqrt{5}, 2\sqrt{35}$.

(k) The basis is

$$\left\{ \left(-\frac{4}{\sqrt{47}}, \frac{3-2i}{\sqrt{47}}, \frac{i}{\sqrt{47}}, \frac{1-4i}{\sqrt{47}} \right), \left(\frac{3-i}{2\sqrt{15}}, -\frac{5i}{2\sqrt{15}}, \frac{2i-1}{\sqrt{15}}, \frac{i+2}{2\sqrt{15}} \right), \right.$$

$$\left. \left(\frac{-i-17}{2\sqrt{290}}, \frac{8i-9}{2\sqrt{290}}, \frac{8i-9}{\sqrt{290}}, \frac{8i-9}{2\sqrt{290}} \right) \right\}.$$

And the Fourier coefficients are $-\sqrt{47}i - \sqrt{47}, 4\sqrt{15}i - 2\sqrt{15}, 2\sqrt{290}i + 2\sqrt{290}$.

(l) The basis is

$$\left\{ \begin{pmatrix} \frac{1-i}{2\sqrt{10}} & \frac{-3i-2}{2\sqrt{10}} \\ \frac{2i+2}{2\sqrt{10}} & \frac{i+4}{2\sqrt{10}} \end{pmatrix}, \begin{pmatrix} \frac{3\sqrt{2}i}{5} & \frac{-i-1}{5\sqrt{2}} \\ \frac{1-3i}{5\sqrt{2}} & \frac{i+1}{5\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{-43i-2}{5\sqrt{323}} & \frac{1-21i}{5\sqrt{323}} \\ -\frac{68i}{5\sqrt{323}} & \frac{34i}{5\sqrt{323}} \end{pmatrix} \right\}.$$

And the Fourier coefficients are $2\sqrt{10} - 6\sqrt{10}i, 10\sqrt{2}, 0$.

(m) The basis is

$$\left\{\begin{pmatrix} \frac{i-1}{3\sqrt{2}} & -\frac{i}{3\sqrt{2}} \\ \frac{2-i}{3\sqrt{2}} & \frac{3i+1}{3\sqrt{2}} \end{pmatrix}, \begin{pmatrix} -\frac{4i}{\sqrt{246}} & \frac{-9i-11}{\sqrt{246}} \\ \frac{5i+1}{\sqrt{246}} & \frac{1-i}{\sqrt{246}} \end{pmatrix}, \begin{pmatrix} \frac{-118i-5}{\sqrt{39063}} & \frac{-26i-7}{\sqrt{39063}} \\ -\frac{145i}{\sqrt{39063}} & -\frac{58}{\sqrt{39063}} \end{pmatrix}\right\}.$$

And the Fourier coefficients are $3\sqrt{2}i + 6\sqrt{2}, -\sqrt{246}i - \sqrt{246}, 0$.

3. Check that β is an orthonormal basis. So we have the coefficients of $(3,4)$ are

$$(3,4) \cdot (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) = \frac{7}{\sqrt{2}}$$

and

$$(3,4) \cdot (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}) = -\frac{1}{\sqrt{2}}.$$

4. We may find the null space for the following system of equations

$$(a,b,c) \cdot (1,0,i) = a - ci = 0,$$

$$(a,b,c) \cdot (1,2,1) = a + 2b + c = 0.$$

So the solution set is $S^\perp = \text{span}\{(i, -\frac{1}{2}(1+i), 1)\}$.

5. We may thought x_0 as a direction, and thus $S_0^\perp$ is the plane orthogonal to it. Also, we may though the span of x_1, x_2 is a plane, and thus $S_0^\perp$ is the line orthogonal to the plane.

6. Take X to be the space generated by W and $\{x\}$. Thus X is also a finite-dimensional subspace. Apply Theorem 6.6 to x in X. We know that x could be uniquely written as $u + v$ with $u \in W$ and $v \in V^\perp$. Since $x \notin W$, we have $v \neq 0$. Pick $y = v$. And we have

$$\langle x, y \rangle = \langle v, v \rangle + \langle u, v \rangle = \|v\|^2 > 0.$$

7. The necessity comes from the definition of orthogonal complement, since every element in β is an element in W. For the sufficiency, assume that $\langle z, v \rangle = 0$ for all $v \in \beta$. Since β is a basis, every element in W could be written as

$$\sum_{i=1}^{k} a_i v_i,$$

where a_i is some scalar and v_i is element in β. So we have

$$\langle z, \sum_{i=1}^{k} a_i v_i \rangle = \sum_{i=1}^{k} \overline{a_i} \langle z, v_i \rangle = 0.$$

Hence z is an element in $W^\perp$.

8. We apply induction on n. When $n = 1$, the Gram-Schmidt process always preserve the first vector. Suppose the statement holds for $n \leq k$. Consider the a orthogonal set of nonzero vectors

$$\{w_1, w_2, \ldots, w_k\}.$$

By induction hypothesis, we know that the vectors $v_i = w_i$ for $i = 1, 2, \ldots, k-1$, where v_i is the vector derived from the process. Now we apply the process the find

$$v_n = w_n - \sum_{i=1}^{k-1} \frac{\langle w_n, v_i \rangle}{\|v_i\|^2} v_i$$

$$= w_n - 0 = w_n.$$

So we get the desired result.

9. The orthonormal basis for W is the set consisting of the normalized vector $(i, 0, 1)$, $\{\frac{1}{\sqrt{2}}(i, 0, 1)\}$. To find a basis for $W^\perp$ is to find a basis for the null space of the following system of equations

$$(a, b, c) \cdot (i, 0, 1) = -ai + c = 0.$$

The basis would be $\{(1, 0, i), (0, 1, 0)\}$. It's lucky that it's orthogonal. If it's not, we should apply the Gram-Schmidt process to it. Now we get the orthonormal basis

$$\{\frac{1}{\sqrt{2}}(1, 0, i), (0, 1, 0)\}$$

by normalizing those elements in it.

10. By Theorem 6.6, we know that $V = W \oplus W^\perp$ since $W \cap W^\perp = \{0\}$ by definition. So there's a nature projection T on W along $W^\perp$. That is, we know tht every element x in V could be writen as $u + v$ such that $u \in W$ and $v \in W^\perp$ and we define $T(x) = u$. Naturally, the null space $N(T)$ is $W^\perp$. And since u and v is always orthogonal, we have

$$\|x\|^2 = \|u\|^2 + \|v\|^2 \geq \|u\|^2 = \|T(x)\|^2$$

by Exercise 6.1.10. And so we have

$$\|T(x)\| \leq \|x\|.$$

11. Use the fact

$$(AA^*)_{ij} = \langle v_i, v_j \rangle$$

for all i and j, where v_i is the i-th row vector of A.

12. If $x \in (R(L_{A^*}))^\perp$, this means that x is orthogonal to $A * y$ for all $y \in \mathbb{F}^m$. So we have

$$0 = \langle x, A^* y \rangle = \langle Ax, y \rangle$$

for all y and hence $Ax = 0$ and $x \in N(L_A)$. Conversely, if $x \in N(L_A)$, we have $Ax = 0$. And

$$\langle x, A * y \rangle = \langle Ax, y \rangle = 0$$

for all y. So x is a element in $(R(L_{A*}))^\perp$.

13. (a) If $x \in S^\perp$ then we have that x is orthogonal to all elements of S, so are all elements of S_0. Hence we have $x \in S_0^\perp$.

(b) If $x \in S$, we have that x is orthogonal to all elements of $S^\perp$. This means x is also an element in $(S^\perp)^\perp$. And

$$\text{span}(S) \subset (S^\perp)^\perp$$

is because span(S) is the smallest subspace containing S and every orthogonal complement is a subspace.

(c) By the previous argument, we already have that $W \subset (W^\perp)^\perp$. For the converse, if $x \notin W$, we may find $y \in W^\perp$ and $\langle x, y \rangle \neq 0$. This means that $W \supset (W^\perp)^\perp$.

(d) By Theorem 6.6, we know that $W = W + W^\perp$. And if $x \in W \cap W^\perp$, we have

$$\langle x, x \rangle = \|x\|^2 = 0.$$

Combine these two and get the desired conclusion.

14. We prove the first equality first. If $x \in (W_1 + W_2)^\perp$, we have x is orthogonal to $u + v$ for all $u \in W_1$ and $v \in W_2$. This means that x is orthogonal to $u = u + 0$ for all u and so x is an element in $W_1^\perp$. Similarly, x is also an element in $W_2^\perp$. So we have

$$(W_1 + W_2)^\perp \subset W_1^\perp.$$

Conversely, if $x \in W_1^\perp \cap W_2^\perp$, then we have

$$\langle x, u \rangle = \langle x, v \rangle = 0$$

for all $u \in W_1$ and $v \in W_2$. This means

$$\langle x, u + v \rangle = \langle x, u \rangle + \langle x, v \rangle = 0$$

for all element $u + v \in W_1 + W_2$. And so

$$(W_1 + W_2)^\perp \supset W_1^\perp.$$

For the second equality, we have ,by Exercise 6.2.13(c),

$$(W_1 \cap W_2)^\perp = ((W_1^\perp)^\perp \cap (W_2^\perp)^\perp)^\perp$$

$$= ((W_1^\perp + W_2^\perp)^\perp)^\perp = W_1^\perp + W_2^\perp.$$

166

15. (a) By Theorem 6.5, we have

$$x = \sum_{i=1}^{n} \langle x, v_i \rangle v_i, y = \sum_{j=1}^{n} \langle y, v_j \rangle v_j.$$

Thus we have

$$\langle x, y \rangle = \sum_{i,j} \langle \langle x, v_i \rangle v_i, \langle y, v_j \rangle v_j \rangle$$

$$= \sum_{i=1}^{n} \langle \langle x, v_i \rangle v_i, \langle y, v_i \rangle v_i \rangle$$

$$= \sum i = 1^n \langle x, v_i \rangle \overline{\langle y, v_i \rangle}.$$

(b) The right hand side of the previous equality is the definition of the standard inner product of $\mathbb{F}^n$.

16. (a) Let $W = \mathrm{span}(S)$. If u is an element in W, who is finite-dimensional, then by Exercise 6.2.15(a) we know that

$$\|u\|^2 = \sum_{i=1}^{n} |\langle u, v_i \rangle|^2.$$

Now for a fixed x, we know that $W' = \mathrm{span}(W \cup \{x\})$ is finite-dimensional. Applying Exercise 6.2.10, we have $T(x) \in W$ and $\|T(x)\| \le \|x\|$. This means

$$\|x\|^2 \ge \|T(x)\|^2 = \sum_{i=1}^{n} |\langle T(x), v_i \rangle|^2$$

by our discussion above. Ultimately, by the definition of T, we have $x = T(x) + y$ for some y who is orthogonal to all the elements in W. Thus we have

$$\langle x, v_i \rangle = \langle T(x), v_i \rangle + \langle y, v_i \rangle = \langle T(x), v_i \rangle.$$

So the inequality holds.

(b) We've explained it.

17. First, since $\langle T(x), y \rangle = 0$ for all y, we have $T(x) = 0$. Applying this argument to all x we get $T(x) = 0$ for all x. For the second version, by Exercise 6.1.9 we know that $T(x) = 0$ for all x in some basis for V. And this means $T = T_0$ by Theorem 2.6.

18. Let f be an odd function. Then for every even funcion g we have fg is an odd function since

$$fg(t) = f(-t)g(-t) = -f(t)g(t).$$

So the inner product of f and g is zero. This means $W_e^{\perp} \supset W_o$.

Conversely, for every function h, we could write $h = f + g$, where

$$f(t) = \frac{1}{2}(h(t) + h(-t))$$

and

$$g(t) = \frac{1}{2}(h(t) - h(-t)).$$

If now h is an element in $W_e^\perp$, we have

$$0 = \langle h, f \rangle = \langle f, f \rangle + \langle g, f \rangle = \|f\|^2$$

since f is a even function. This means that $f = 0$ and $h = g$, an element in W_o.

19. Find an orthonormal basis for W and use the formula in Theorem 6.6.

(a) Pick $\{\frac{1}{\sqrt{17}}(1, 4)\}$ as a basis for W. Thus the orthogonal projection of u is

$$\langle u, \frac{1}{\sqrt{17}}(1, 4) \rangle \frac{1}{\sqrt{17}}(1, 4) = \frac{26}{17}(1, 4).$$

(b) Pick $\{\frac{1}{\sqrt{10}}(-3, 1, 0), \frac{1}{\sqrt{5}}(2, 1, 0)\}$ as a basis for W. Thus the orthogonal projection of u is

$$\langle u, \frac{1}{\sqrt{10}}(-3, 1, 0) \rangle \frac{1}{\sqrt{10}}(-3, 1, 0) + \langle, \frac{1}{\sqrt{35}}(1, 3, 5) \rangle \frac{1}{\sqrt{35}}(1, 3, 5)$$

$$= -\frac{1}{2}(-3, 1, 0) + \frac{4}{7}(1, 3, 5) = \frac{1}{14}(29, 17, 40).$$

(c) Pick $\{1, \frac{1}{\sqrt{3}}(2x-1)\}$ as a basis for W. Thus the orthogonal projection of h is

$$\langle h, 1 \rangle 1 + \langle h, \frac{1}{\sqrt{3}}(2x - 1) \rangle \frac{1}{\sqrt{3}}(2x - 1)$$

$$= \frac{29}{6} + \frac{1}{18}(2x - 1) = \frac{1}{9}(x + 1).$$

20. If v is the orthogonal projection of u, then the distance of u is the length of $u - v$.

(a) The distance is

$$\|(2, 6) - \frac{26}{17}(1, 4)\| = \frac{2}{\sqrt{17}}.$$

(b) The distance is

$$\|(2, 1, 3) - \frac{1}{14}(29, 17, 40)\| = \frac{1}{\sqrt{14}}.$$

(c) The distance is

$$\|(-2x^2 + 3x + 4) - \frac{1}{9}(x + 1)\| = \sqrt{\frac{30527}{1215}}.$$

21. Do this with the same method as that in Exercise 6.2.19. Let $\{u_1, u_2, u_3\}$ be the orthonormal basis given by Example 5 of this section. Then the closest second-degree polynomial approximation is the orthogonal projection

$$\langle e^t, u_1\rangle u_1 + \langle e^t, u_2\rangle u_2 + \langle e^t, u_3\rangle u_3$$

$$= \frac{1}{4}[(15e - 105e^{-1})t^2 + 12t - 3e^2 + 33].$$

22. (a) Use the Gram-Schmidt process to find a orthogonal basis $\{t, -\frac{6t-5\sqrt{t}}{5}\}$ and normalize it to be

$$\{\sqrt{3}t, -\sqrt{2}(6t - 5\sqrt{t})\}.$$

(b) Do it as that we've done in Exercise 6.2.19. The approximation is

$$-\frac{20\sqrt{t} - 45t}{28}.$$

23. (a) Let $x(n), y(n), z(n)$ be sequences in the condition of inner product space. Since all of them has entry not zero in only finite number of terms, we may find an integer N such that

$$x(n) = y(n) = z(n)$$

for all $n \geq N$. But this means that all of them are vectors in $\mathbb{F}^N$. So it's an inner product.

(b) It's orthogonal since

$$\langle e_i, e_j\rangle = \sum_{n=1}^{\infty} e_i(n)\overline{e_j(n)}$$

$$= e_i(i)\overline{e_j(i)} + e_i(j)\overline{e_j(j)} = 0.$$

And it's orthonormal since

$$\langle e_i, e_i\rangle = \sum_{n=1}^{\infty} e_i(n)\overline{e_i(n)} = e_i(i)\overline{e_i(i)} = 1.$$

(c) i. If e_1 is an element in W, we may write

$$e_1 = a_1\sigma_1 + a_2\sigma_2 + \cdots + a_k\sigma_k,$$

where a_i is some scalar. But we may observe that a_i must be zero otherwise the i-th entry of e_1 is nonzero, which is impossible. So this means that $e_1 = 0$. It's also impossible. Hence e_1 cannot be an element in W.

ii. If a is a sequence in $W^\perp$, we have $a(1) = -a(n)$ for all i since

$$\langle a, \sigma_n \rangle = a(1) + a(n) = 0$$

for all i. This means that if a contains one nonzero entry, then all entries of a are nonzero. This is impossible by our definition of the space V. Hence the only element in $W^\perp$ is zero.

On the other hand, we have $W^\perp = \{0\}^\perp = V$. But by the previous argument, we know that $W \neq V = (W^\perp)^\perp$.

6.3 The Adjoint of a Linear Operator

1. (a) Yes. See Theorem 6.9.

   (b) No. It just for the linear mapping $V \to \mathbb{F}$. For example, the value of the identity mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$ is not an element in $\mathbb{F}$.

   (c) No. The equality holds only for the case β is an orthonormal basis. For example, let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $T = L_A$. Thus we have $T^* = L_{A^*}$. But for the basis
   $$\beta = \{(1,1), (0,1)\},$$
   we have
   $$[T]_\beta = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} \neq [T^*]_\beta = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

   (d) Yes.See Theorem 6.9.

   (e) No. Choose $a = i$, $b = 0$ and $T = I_\mathbb{C} = U$. We have $(aT)^* = \bar{a}T^* \neq aT^*$.

   (f) Yes. See Theorem 6.10 and its Corollary.

   (g) Yes. See Theorem 6.11.

2. Follow the prove of Theorem 6.8. The vector would be

$$y = \sum_{i=1}^{n} \overline{g(v_i)} v_i.$$

   (a) The vector is $(1, -2, 4)$.

   (b) The vector is $(1, 2)$.

   (c) The vector is $210x^2 - 204x + 33$.

3. Use the definition and that skill used in the previous exercises.

   (a) By definition we have

$$\langle (a,b), T^*(x) \rangle = \langle (2a + b, a - 3b), (3,5) \rangle = 11a - 12b.$$

   We may observe that $T^*(x) = (11, -12)$.

(b) By definition we have

$$\langle (z_1, z_2), T^*(x) \rangle = \langle (2z_1 + iz_2, (1-i)z_1), (3-i, 1+2i) \rangle$$

$$= (5-i)z_1 + (-1+3i)z_2.$$

We may observe that $T^*(x) = (5+i, -1-3i)$.

(c) By definition we have note inner product invovles conjugate ...

$$\langle at + b, T^*(x) \rangle = \langle 3at + (a+3b), 4-2t \rangle$$

$$= \int_{-1}^{1} -6at^2 + (-2a+6b)t + (4a+12b)dt$$

We may observe that $T^*(x) = 6t + 12$.

4. **(b)** Compute

$$\langle x, (cT)^*(y) \rangle = \langle cT(x), y \rangle =$$

$$\langle T(x), \bar{c}y \rangle = \langle x, T^*(\bar{c}y) \rangle = \langle x, \bar{c}T^*(y) \rangle$$

for all x and y.

(c) Compute

$$\langle x, (TU)^*(y) \rangle = \langle TU(x), y \rangle$$

$$= \langle U(x), T^*(y) \rangle = \langle x, U^*T^*(y) \rangle$$

for all x and y.

(e) Compute

$$\langle x, I^*(y) \rangle = \langle I(x), y \rangle$$

$$= \langle x, y \rangle = \langle x, I(y) \rangle$$

for all x and y.

5. (a) Just write it down just as that in the proof of (c).

(a) Compute

$$L_{(A+B)^*} = (L_{A+B}^* = (L_A + L_B)^*$$

$$= (L_A)^* + (L_B)^* = L_{A^*} + L_{B^*} = L_{A^*+B^*}.$$

(b) Compute

$$L_{(cA)^*} = (L_{cA})^* = (cL_A)^*$$

$$= \bar{c}(L_A)^* = \bar{c}L_{A^*} = L_{\bar{c}A^*}.$$

(d) Compute

$$L_{A^{**}} = (L_A)^{**} = L_A.$$

(e) Compute

$$L_{I^*} = (L_I)^* = L_I.$$

(b) The statement for nonsquare matrices has no difference with the Corollary but the statement (e) cannot holds since there's no non-square identity matrix. To this result come from that A^* is the conjugate of A^t.

6. Compute
$$U_1^* = (T + T^*)^* = T^* + T^{**} = T^* + T = U_1$$
and
$$U_2^* = (TT^*)^* = T^{**}T^* = TT^* = U_2.$$

7. Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$. Then $N(A) \neq N(A^*)$ since $(0,1)$ is an element only in the later one.

8. If T is invertible, then we have the inverse mapping T^{-1}. Then we have
$$T^*(T^{-1})^* = (T^{-1}T)^* = I^* = I.$$

9. For each vector $v \in V$ we may write $v = v_1 + v_2$ such that $v_1 \in W$ and $v_2 \in W^\perp$. Now check
$$\langle x_1 + x_2, T^*(y_1 + y_2) \rangle = \langle T(x_1 + x_2), y_1 + y_2 \rangle$$
$$= \langle x_1, y_1 + y_2 \rangle = \langle x_1, y_1 \rangle$$
and
$$\langle x_1 + x_2, T(y_1 + y_2) \rangle = \langle x_1 + x_2, y_1 \rangle$$
$$= \langle x_1, y_1 \rangle = \langle x_1 + x_2, T^*(y_1 + y_2) \rangle$$
for all $x = x_1 + x_2$ and $y = y_1 + y_2$.

10. The sufficiency is easy since we may just pick $y = x$. For the necessity, suppose now $\|T(x)\| = \|x\|$. By Exercise 6.1.20 we have
$$\langle x, y \rangle = \frac{1}{4} \sum_{k=1}^4 i^k \|x + i^k y\|^2$$
$$= \frac{1}{4} \sum_{k=1}^4 i^k \|T(x + i^k y)\|^2 = \frac{1}{4} \sum_{k=1}^4 i^k \|T(x) + i^k T(y)\|^2$$
$$\langle T(x), T(y) \rangle$$

if $\mathbb{F} = \mathbb{C}$. However, for the case $\mathbb{F} = \mathbb{R}$ the above argument could also work if we just pick k to be 2 and 4.

11. We have
$$0 = \langle T^*T(x), x \rangle = \langle T(x), T(x) \rangle = \|T(x)\|^2$$

for all x and hence $T(x) = 0$ for all x. And the second statement is also true since we may write

$$TT^* = T^{**}T^* = T_0$$

and get $T^* = T_0$. Since $T = T^{**} = T_0^* = T_0$.

12. (a) If $x \in R(T^*)^\perp$ we have

$$0 = \langle x, T^*(y) \rangle = \langle T(x), y \rangle$$

for all y. This means that $T(x) = 0$ and so $x \in N(T)$. Conversely, if $x \in N(T)$, we have

$$\langle x, T^*(y) \rangle = \langle T(x), y \rangle = 0$$

for all y. This means that x is an element in $R(T^*)^\perp$.

(b) By Exercise 6.2.13(c) we have

$$N(T)^\perp = (R(T^*)^\perp)^\perp = R(T^*)^\perp.$$

13. (a) If $x \in N(T^*T)$ we have $T^*T(x) = 0$ and

$$0 = \langle T^*T(x), x \rangle = \langle T(x), T(x) \rangle.$$

This means that $T(x) = 0$ and $x \in N(T)$. Conversely, if $x \in N(T)$, we have $T^*T(x) = T^*(0) = 0$ and so $x \in N(T^*T)$.

On the other hand, since the dimension is finite, we have

$$R(T^*T) = N(T^*T)^\perp = N(T)^\perp = R(T^*)$$

by the previous exercise. Hence we have

$$\text{rank}(T^*T) = \text{rank}(T^*) = \text{rank}(T)$$

by the next argument.

(b) For arbitrary matrix A, denote $\overline{A}$ to be the matrix consisting of the conjugate of entris of A. Thus we have $A^* = \overline{A^t}$. We want to claim that $\text{rank}(A) = \text{rank}(A^*)$ first. Since we already have that $\text{rank}(A) = \text{rank}(A^t)$, it's sufficient to show that $\text{rank}(A) = \text{rank}(\overline{A})$. By Theorem 3.6 and its Corollaries, we may just prove that $\{v_i\}_{i \in I}$ is independent if and only if $\{\overline{v_i}\}_{i \in I}$ is independent, where $\overline{v_i}$ means the vector obtained from v_i by taking conjugate to each coordinate. And it comes from the fact

$$\sum_{i \in I} a_i v_i = 0$$

173

if and only if

$$\sum_{i \in I} \overline{a_i v_i} = \overline{\sum_{i \in I} a_i v_i} = 0.$$

Finally, by Theorem 6.10 we already know that $[T]_\beta^* = [T^*]_\beta$ for some basis β. This means that $\text{rank}(T) = \text{rank}(T^*)$. And so

$$\text{rank}(TT^*) = \text{rank}(T^{**}T^*) = \text{rank}(T^*) = \text{rank}(T).$$

    (c) It comes from the fact $L_{A^*} = (L_A)^*$.

14. It's linear since

$$T(cx_1 + x_2) = \langle cx_1 + x_2, y \rangle z$$

$$= c\langle x_1, y \rangle z + \langle x_2, y \rangle z = cT(x_1) + T(x_2).$$

On the other hand, we have

$$\langle u, T^*(v) \rangle = \langle \langle u, y \rangle z, v \rangle$$

$$= \langle u, y \rangle \langle z, v \rangle = \langle u, \langle v, z \rangle y \rangle$$

for all u and v. So we have

$$T^*(x) = \langle x, z \rangle y.$$

15.   (a) Let $y \in W$ be given. We may define

$$g_y(x) = \langle T(x), y \rangle_2,$$

which is linear since T is linear and the first component of inner product function is also linear. By Theorem 6.8 we may find an unique vector, called $T^*(y)$, such that

$$\langle x, T^*(y) \rangle_1 = \langle T(x), y \rangle_2$$

for all x. This means $T^*(y)$ is always well-defined. It's unique since

$$\langle x, T^*(y) \rangle_1 = \langle x, U(y) \rangle_1$$

for all x and y implies that $T^* = U$.
Finally, it's also linear since

$$\langle x, T^*(y + cz) \rangle_1 = \langle T(x), y + cz \rangle_2$$

$$= \langle T(x), y \rangle_2 + \overline{c} \langle T(x), z \rangle_2$$

$$= \langle x, T^*(y) \rangle_1 + \overline{c} \langle x, T^*(z) \rangle_1$$

$$= \langle x, T^*(y) \rangle_1 + \langle x, cT^*(z) \rangle_1 = \langle x, T^*(y) + cT^*(z) \langle_1$$

for all x, y, and z.

(b) Let
$$\beta = \{v_1, v_2, \ldots, v_m\}$$
and
$$\gamma = \{u_1, u_2, \ldots, u_n\}.$$
Further, assume that
$$T(v_j) = \sum_{i=1}^{n} a_{ij} u_i.$$
This means that $[T]_\beta^\gamma = \{a_{ij}\}$.
On the other hand, assume that
$$T^*(u_j) = \sum_{i=1}^{n} c_{ij} v_i.$$

And this means
$$\overline{c_{ij}} = \langle v_i, T^*(u_j)\rangle_1 = \langle T(v_i), u_j\rangle_2 = a_{ji}$$
and $[T^*]_\gamma^\beta = ([T]_\beta^\gamma)^*$.

(c) It comes from the same reason as Exercise 6.3.13(b).

(d) See
$$\langle T^*(x), y\rangle = \overline{\langle y, T^*(x)\rangle}$$
$$= \overline{\langle T(y), x\rangle} = \langle x, T^*(y)\rangle.$$

(e) If $T(x) = 0$ we have $T^*T(x) = T^*(0) = 0$. If $T^*T(x) = 0$ we have
$$0 = \langle x, T^*T(x)\rangle = \langle T(x), T(x)\rangle$$
and henve $T(x) = 0$.

16. (a) Compute
$$\langle x, (T+U)^*(y)\rangle_1 = \langle (T+U)(x), y\rangle_2$$
$$= \langle T(x) + U(x), y\rangle_2 = \langle T(x), y\rangle_2 + \langle U(x), y\rangle_2$$
$$= \langle x, T^*(y)\rangle_1 + \langle x, U^*(y)\rangle_1 = \langle x, (T^* + U^*)(y)\rangle_1$$
for all x and y.

(b) Compute
$$\langle x, (cT)^*(y)\rangle_1 = \langle cT(x), y\rangle_2 =$$
$$\langle T(x), \bar{c}y\rangle_2 = \langle x, T^*(\bar{c}y)\rangle_1 = \langle x, \bar{c}T^*(y)\rangle_1$$
for all x and y.

(c) Let T is a mapping on W and U is a mapping from V to W. Compute

$$\langle x, (TU)^*(y)\rangle_1 = \langle TU(x), y\rangle_2$$

$$= \langle U(x), T^*(y)\rangle_2 = \langle x, U^*T^*(y)\rangle_1$$

for all x and y.

Let T is a mapping from V to W and U is a mapping on V. Compute

$$\langle x, (TU)^*(y)\rangle_1 = \langle TU(x), y\rangle_2$$

$$= \langle U(x), T^*(y)\rangle_1 = \langle x, U^*T^*(y)\rangle_1$$

for all x and y.

(d) Compute

$$\langle x, T^{**}(y)\rangle_2 = \langle T^*(x), y\rangle_1$$

$$= \langle x, T(y)\rangle$$

for all x and y.

17. If $x \in R(T^*)^\perp$ we have

$$0 = \langle x, T^*(y)\rangle_1 = \langle T(x), y\rangle_2$$

for all y. This means that $T(x) = 0$ and so $x \in N(T)$. Conversely, if $x \in N(T)$, we have

$$\langle x, T^*(y)\rangle_1 = \langle T(x), y\rangle_2 = 0$$

for all y. This means that x is an element in $R(T^*)^\perp$.

18. For arbitrary matrix M we already have $\det(M) = \det(M^t)$. So it's sufficient to show that $\overline{\det(M)} = \det(\overline{M})$. We prove this by induction on n, the size of a matrix. For $n = 1$, we have

$$\overline{\det\begin{pmatrix} a \end{pmatrix}} = \det\begin{pmatrix} \bar{a} \end{pmatrix}.$$

For $n = 2$, we have

$$\overline{\det\begin{pmatrix} a & b \\ c & d \end{pmatrix}} = \overline{ad - bc}$$

$$= \bar{a}\bar{d} - \bar{b}\bar{c} = \det\begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix}.$$

Suppose the hypothesis is true for $n = k - 1$. Consider a $k \times k$ matrix M. We have

$$\overline{\det(M)} = \overline{\sum_{j=1}^{k} (-1)^{i+j} M_{ij} \cdot \det(\tilde{M}_{ij})}$$

$$= \sum_{j=1}^{k} (-1)^{i+j} \overline{M_{ij}} \cdot \overline{\det(\tilde{M}_{ij})}$$

176

$$= \sum_{j=1}^{k} (-1)^{i+j} \overline{M_{ij}} \cdot \det(\widetilde{\overline{M}}_{ij}) = \det(\overline{M}).$$

This means that

$$\det(A) = \det(A^t) = \det(A^*).$$

19. Let v_i be the i-th column of A. Then we have v_i^* is the i-th row of A^*. And the desired result comes from the fact

$$(A^*A)_{ij} = v_i^* v_j = \langle v_j, v_i \rangle,$$

which is zero when $i \neq j$.

20. Follow the method after Theorem 6.12.

   (a) The linear function is $-2t + \frac{5}{2}$ with error $E = 1$. The quadratic function is $\frac{1}{3}t^2 - \frac{4}{3}t + 2$ with the error $E = 0$.

   (b) The linear function is $\frac{5}{4}t + \frac{11}{20}$ with error $E = \frac{3}{10}$. The quadratic function is $\frac{1}{56}t^2 + \frac{15}{14} + \frac{239}{280}$ with the error $E = \frac{8}{35}$.

   (c) The linear function is $-\frac{9}{5}t + \frac{4}{5}$ with error $E = \frac{2}{5}$. The quadratic function is $-\frac{1}{7}t^2 - \frac{9}{5}t + \frac{38}{35}$ with the error $E = \frac{4}{35}$.

21. Follow the same method. We have the linear function is $2.1x - \frac{127}{20}$. So the spring constant is 2.1.

22. As the statement in Theorem 6.13, we may first find a vector u such that $AA^*u = b$. Finally the minimal solution would be A^*u.

   (a) The minimal solution is $(2, 4, -2)$.

   (b) The minimal solution is $\frac{1}{7}(2, 3, 1)$.

   (c) The minimal solution is $(1, , -\frac{1}{2}, \frac{1}{2})$.

   (d) The minimal solution is $\frac{1}{12}(7, 1, 3, -1)$.

23. (a) Direct calculate that

$$A^*A = \begin{pmatrix} \sum_{i=1}^{m} t_i^2 & \sum_{i=1}^{m} t_i \\ \sum_{i=1}^{m} t_i & m \end{pmatrix}$$

and get the result by

$$A^*A \begin{pmatrix} c \\ d \end{pmatrix} = A^*y.$$

For the second method, we may calculate that

$$E = \sum_{i=1}^{m} (y_i - ct_i - d)^2$$

and $\frac{\partial E}{\partial c} = 0$ and $\frac{\partial E}{\partial d} = 0$ give the normal equations.

(b) We want to claim that $\bar{y} = c\bar{t}+d$, whrer $\bar{t}$, $\bar{y}$ are defined in the question and c, d are a solution of the normal equations. But this is an instant result by dividing the second equation by m.

24. (a) Check

$$T(c\sigma + \tau)(k) = \sum_{i=k}^{\infty}(c\sigma + \tau)(k)$$

$$= c\sum_{i=k}^{\infty}(\sigma)(k) + \sum_{i=k}^{\infty}(\tau)(k)$$

$$= cT(\sigma)(k) + T(\tau)(k).$$

(b) For $k \le n$ we have

$$T(e_n)(k) = \sum_{i=k}^{\infty}e_n(i) = 1 = \sum_{i=1}^{n}e_i(k).$$

And for $k > n$ we have

$$T(e_n)(k) = \sum_{i=k}^{\infty}e_n(i) = 0 = \sum_{i=1}^{n}e_i(k).$$

(c) Suppoe that T^* exist. We try to compute $T^*(e_1)$ by

$$1 \cdot \overline{T^*(e_1)}(i) = \langle e_i, T^*(e_1)\rangle = \langle\sum_{i=1}^{n}e_i, e_1\rangle = 1.$$

This means that $T^*(e_1)(i) = 1$ for all i. This is impossible since $T^*(e_1)$ is not an element in V.

6.4   Normal and Self-Adjoint Operators

1. (a) Yes. Check $TT^* = T^2 = T^*T$.

(b) No. The two matrices $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ have $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to be their unique normalized eigenvectors respectly.

(c) No. Consider $T(a,b) = (2a, b)$ to be a mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$ and β to be the basis $\{(1,1),(1,0)\}$. We have T is normal with $T^* = T$. But $[T]_\beta = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$ is not normal. Furthermore, the converse is also not true. We may let $T(a,b), = (b,b)$ be a mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$ and β be the basis $\{(1,-1),(0,1)\}$. In this time T is not normal with $T^*(a,b) = (0, a+b)$. However, $[T]_\beta = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ is a normal matrix.

(d) Yes. This comes from Theorem 6.10.

(e) Yes. See the Lemma before Theorem 6.17.

(f) Yes. We have $I^* = I$ and $O^* = O$, where I and O are the identity and zero operators.

(g) No. The mapping $T(a, b) = (-b, a)$ is normal since $T^*(a, b) = (a, -b)$. But it's not diagonalizable since the characteristic polynomial of T does not split.

(h) Yes. If it's an operator on a real inner product space, use Theorem 6.17. If it's an operator on a complex inner product space, use Theorem 6.16.

2. Use one orthonormal basis β to check $[T]_\beta$ is normal, self-adjoint, or neither. Ususally we'll take β to be the standard basis. To find an orthonormal basis of eigenvectors of T for V, just find an orthonormal basis for each eigenspace and take the union of them as the desired basis.

(a) Pick β to be the standard basis and get that

$$[T]_\beta = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}.$$

So it's self-adjoint. And the basis is

$$\{\frac{1}{\sqrt{5}}(1, -2), \frac{1}{\sqrt{5}}(2, 1)\}.$$

(b) Pick β to be the standard basis and get that

$$[T]_\beta = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 5 & 0 \\ 4 & -2 & 5 \end{pmatrix}.$$

So it's neither normal nor self-adjoint.

(c) Pick β to be the standard basis and get that

$$[T]_\beta = \begin{pmatrix} 2 & i \\ 1 & 2 \end{pmatrix}.$$

So it's normal but not self-adjoint. And the basis is

$$\{(\frac{1}{\sqrt{2}}, -\frac{1}{2} + \frac{1}{2}i), (\frac{1}{\sqrt{2}}, \frac{1}{2} - \frac{1}{2}i)\}.$$

(d) Pick an orthonormal basis $\beta = \{1, \sqrt{3}(2t - 1), \sqrt{6}(6t^2 - 6t + 1)\}$ by Exercise 6.2.2(c) and get that

$$[T]_\beta = \begin{pmatrix} 0 & 2\sqrt{3} & 0 \\ 0 & 0 & 6\sqrt{2} \\ 0 & 0 & 0 \end{pmatrix}.$$

So it's neither normal nor self-adjoint.

(e) Pick β to be the standard basis and get that

$$[T]_\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

So it's self-adjoint. And the basis is

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

(f) Pick β to be the standard basis and get that

$$[T]_\beta = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

So it's self-adjoint. And the basis is

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

3. Just see Exercise 1(c).

4. Use the fact

$$(TU)^* = U^*T^* = UT.$$

5. Observe that $(T - cI)^* = T^* - \bar{c}I$ and check

$$(T - cI)(T - cI)^* = (T - cI)(T^* - \bar{c}I) = TT^* - \bar{c}T - cT^* + |c|^2 I$$

and

$$(T - cI)^*(T - cI) = (T^* - \bar{c}I)(T - cI) = T^*T - \bar{c}T - cT^* + |c|^2 I.$$

They are the same because $TT^* = T^*T$.

6. (a) Observe the fact

$$T_1^* = (\frac{1}{2}(T + T^*))^* = \frac{1}{2}(T^* + T) = T_1$$

and

$$T_2^* = (\frac{1}{2i}(T - T^*))^* = -\frac{1}{2i}(T^* - T) = T_2.$$

(b) Observe that $T^* = U_1^* - iU_2^* = U_1 - iU_2$. This means that

$$U_1 = \frac{1}{2}(T + T^*) = T_1$$

and

$$U_2 - \frac{1}{2}(T - T^*) = T_2.$$

180

(c) Calculate that

$$T_1 T_2 - T_2 T_1 = \frac{1}{4i}(T^2 - TT^* + T^*T - (T^*)^2) - \frac{1}{4i}(T^2 + TT^* - T^*T - (T^*)^2)$$

$$= \frac{1}{2i}(T^*T - TT^*).$$

It equals to T_0 if and only if T is normal.

7. (a) We check
$$\langle x, (T_W)^*(y) \rangle = \langle T_W(x), y \rangle = \langle T(x), y \rangle$$
$$= \langle T^*(x), y \rangle = \langle x, T(y) \rangle = \langle x, T_W(y) \rangle$$
for all x and y in W.

(b) Let y be an element in $W^\perp$. We check

$$\langle x, T^*(y) \rangle = \langle T(x), y \rangle = 0$$

for all $x \in W$, since $T(x)$ is also an element in W by the fact that W is T-invariant.

(c) We check
$$\langle x, (T_W)^*(y) \rangle = \langle T_W(x), y \rangle = \langle T(x), y \rangle$$
$$\langle x, T^*(y) \rangle = \langle x, (T^*)_W(y) \rangle.$$

(d) Since T is normal, we have $TT^* = T^*T$. Also, since W is both T- and T^*-invariant, we have

$$(T_W)^* = (T^*)_W$$

by the previous argument. This means that

$$T_W(T_W)^* = T_W(T^*)_W = (T^*)_W T_W = (T_W)^* T_W.$$

8. By Theorem 6.16 we know that T is diagonalizable. Also, by Exercise 5.4.24 we know that T_W is also diagonalizable. This means that there's a basis for W consisting of eigenvectors of T. If x is a eigenvectors of T, then x is also a eigenvector of T^* since T is normal. This means that there's a basis for W consisting of eigenvectors of T^*. So W is also T-invariant.

9. By Theorem 6.15(a) we know that $T(x) = 0$ if and only if $T^*(x) = 0$. So we get that $N(T) = N(T^*)$. Also, by Exercise 6.3.12 we know that

$$R(T) = N(T^*)^\perp = N(T)^\perp = R(T^*).$$

10. Directly calculate that

$$\|T(x) \pm ix\|^2 = \langle T(x) \pm ix, T(x) \pm ix \rangle =$$

$$= \|T(x)\|^2 \pm \langle T(x), ix \rangle \pm \langle ix, T(x) \rangle + \|x\|^2$$

$$= \|T(x)\|^2 \mp i\langle T(x), x \rangle \pm \langle T^*(x), x \rangle + \|x\|^2 = \|T(x)\|^2 + \|x\|^2.$$

Also, $T \pm iI$ is injective since $\|T(x) \pm x\| = 0$ if and only if $T(x) = 0$ and $x = 0$. Now $T - iI$ is invertible by the fact that V is finite-dimensional. Finally we may calculate that

$$\langle x, [(T - iI)^{-1}]^*(T + iI)(y) \rangle = \langle (T - iI)^{-1}(x), (T + iI)(y) \rangle$$

$$= \langle (T - iI)^{-1}(x), (T^* + iI)(y) \rangle = \langle (T - iI)^{-1}(x), (T - iI)^*(y) \rangle$$

$$= \langle (T - iI)(T - iI)^{-1}(x), y \rangle = \langle x, y \rangle$$

for all x and y. So we get the desired equality.

11. (a) We prove it by showing the value is equal to its own conjugate. That is,

$$\overline{\langle T(x), x \rangle} = \langle x, T(x) \rangle$$

$$\langle x, T^*(x) \rangle = \langle T(x), x \rangle.$$

(b) As Hint, we compute

$$0 = \langle T(x + y), x + y \rangle$$

$$= \langle T(x), x \rangle + \langle T(x), y \rangle + \langle T(y), x \rangle + \langle T(y), y \rangle$$

$$= \langle T(x), y \rangle + \langle T(y), x \rangle.$$

That is, we have

$$\langle T(x), y \rangle = -\langle T(y), x \rangle.$$

Also, replace y by iy and get

$$\langle T(x), iy \rangle = -\langle T(iy), x \rangle$$

and hence

$$-i\langle T(x), y \rangle = -i\langle T(y), x \rangle.$$

This can only happen when

$$\langle T(x), y \rangle = 0$$

for all x and y. So T is the zero mapping.

(c) If $\langle T(x), x \rangle$ is real, we have

$$\langle T(x), x \rangle = \langle x, T(x) \rangle = \langle T^*(x), x \rangle.$$

This means that

$$\langle (T - T^*)(x), x \rangle = 0$$

for all x. By the previous argument we get the desired conclusion $T = T^*$.

12. Since the characteristic polynomial splits, we may apply Schur's Theorem and get an orthonormal basis β such that $[T]_\beta$ is upper triangular. Denote the basis by

$$\beta = \{v_1, v_2, \ldots, v_n\}.$$

We already know that v_1 is an eigenvector. Pick t to be the maximum integer such that $v_1, v_2, \ldots, v_t$ are all eigenvectors with respect to eigenvalues λ_i. If $t = n$ then we've done. If not, we will find some contradiction. We say that $[T]_\beta = \{A_{i,j}\}$. Thus we know that

$$T(v_{t+1}) = \sum_{i=1}^{t+1} A_{i,t+1} v_i.$$

Since the basis is orthonormal, we know that

$$A_{i,t+1} = \langle T(v_{t+1}), v_i \rangle = \langle v_{t+1}, T^*(v_i) \rangle$$

$$= \langle v_{t+1}, \overline{\lambda_i} v_i \rangle = 0$$

by Theorem 6.15(c). This means that v_{t+1} is also an eigenvector. This is a contradiction. So β is an orthonormal basis. By Theorem 6.17 we know that T is self-adjoint.

13. If A is Gramian, we have A is symmetric since $A^t = (B^t B)^t = B^t B = A$. Also, let λ be an eigenvalue with unit eigenvector x. Then we have $Ax = \lambda x$ and

$$\lambda = \langle Ax, x \rangle = \langle B^t Bx, x \rangle = \langle Bx, Bx \rangle \geq 0.$$

Conversely, if A is symmetric, we know that L_A is a self-adjoint operator. So we may find an orthonormal basis β such that $[L_A]_\beta$ is diagonal with the ii-entry to be λi. Denote D to be a diagonal matrix with its ii-entry to be $\sqrt{\lambda_i}$. So we have $D^2 = [L_A]_\beta$ and

$$A = [I]_\beta^\alpha [L_A]_\beta [I]_\alpha^\beta = ([I]_\beta^\alpha D)(D[I]_\alpha^\beta),$$

where α is the standard basis. Since the basis β is orthonormal, we have $[I]_\beta^\alpha = ([I]_\alpha^\beta)^t$. So we find a matrix

$$B = D[I]_\alpha^\beta$$

such that $A = B^t B$.

14. We use induction on the dimension n of V. If $n = 1$, U and T will be diagonalized simultaneously by any orthonormal basis. Suppose the statement is true for $n \leq k - 1$. Consider the case $n = k$. Now pick one arbitrary eigenspace $W = E_\lambda$ of T for some eigenvalue λ. Note that W is T-invariant naturally and U-invariant since

$$TU(w) = UT(w) = \lambda U(w)$$

for all $w \in W$. If $W = V$, then we may apply Theorem 6.17 to the operator U and get an orthonormal basis β consisting of eigenvectors of U. Those vectors will also be eigenvectors of T. If W is a proper subspace of V, we may apply the induction hypothesis to T_W and U_W, which are self-adjoint by Exercise 6.4.7, and get an orthonormal basis β_1 for W consisting of eigenvectors of T_W and U_W. So those vectors are also eigenvectors of T and U. On the other hand, we know that $W^\perp$ is also T- and U-invariant by Exercise 6.4.7. Again, by applying the induction hypothesis we get an orthonormal basis β_2 for $W^\perp$ consisting of eigenvectors of T and U. Since V is finite dimentional, we know that $\beta = \beta_1 \cup \beta_2$ is an orthonormal basis for V consisting of eigenvectors of T and U.

15. Let $T = L_A$ and $U = L_B$. Applying the previous exercise, we find some orthonormal basis β such that $[T]_\beta$ and $[U]_\beta$ are diagonal. Denote α to be the standard basis. Now we have that

$$[T]_\beta = [I]_\alpha^\beta A [I]_\beta^\alpha$$

and

$$[U]_\beta = [I]_\alpha^\beta B [I]_\beta^\alpha$$

are diagonal. Pick $P = [I]_\beta^\alpha$ and get the desired result.

16. By Schur's Theorem $A = P^{-1} B P$ for some upper triangular matrix B and invertible matrix P. Now we want to say that $f(B) = O$ first. Since the characteristic polynomial of A and B are the same, we have the characteristic polynomial of A would be

$$f(t) = \prod_{i=1}^{n} (B_{ii} - t)$$

since B is upper triangular. Let $C = f(B)$ and $\{e_i\}$ the be the standard basis. We have $Ce_1 = 0$ since $(B_{11}I - B)e_1 = 0$. Also, we have $Ce_i = 0$ since $(B_{ii}I - B)e_i$ is a linear combination of $e_1, e_2, \ldots, e_{i-1}$ and so this vector will vanish after multiplying the matrix

$$\prod_{j=1}^{i-1} (B_{ii}I - B).$$

So we get that $f(B) = C = O$. Finally, we have

$$f(A) = f(P^{-1}BP) = P^{-1}f(B)P = O.$$

17. (a) By Theorem 6.16 and Theorem 6.17 we get an orthonormal basis

$$\alpha = \{v_1, v_2, \ldots, v_n\},$$

where v_i is the eigenvector with respect to the eigenvalue λ_i, since T is self-adjoint. For each vector x, we may write it as

$$x = \sum_{i=1}^{n} a_i v_i.$$

Compute

$$\langle T(x), x \rangle = \langle \sum_{i=1}^{n} a_i \lambda_i v_i, \sum_{i=1}^{n} a_i v_i \rangle$$

$$= \sum_{i=1}^{n} |a_i|^2 \lambda_i.$$

The value is greater than [no less than] zero for arbitrary set of a_i's if and only if λ_i is greater than [no less than] zero for all i.

(b) Denote β to be

$$\{e_1, e_2, \ldots, e_n\}.$$

For each $x \in V$, we may write it as

$$x = \sum_{i=1}^{n} a_i e_i.$$

Also compute

$$\langle T(x), x \rangle = \langle \sum_{i=1}^{n} (\sum_{j=1}^{n} A_{ij} a_j) e_i, \sum_{i=1}^{n} a_i e_i \rangle$$

$$= \sum_{i=1}^{n} (\sum_{j=1}^{n} A_{ij} a_j) \overline{a_i} = \sum_{i,j} A_{ij} a_j \overline{a_i}.$$

(c) Since T is self-adjoint, by Theorem 6.16 and 6.17 we have $A = P^* D P$ for some matrix P and some diagonal matrix D. Now if T is positive semidefinite, we have all eigenvalue of T are nonnegative. So the ii-entry of D is nonnegative by the previous argument. We may define a new diagonal matrix E whose ii-entry is $\sqrt{D_{ii}}$. Thus we have $E^2 = D$ and $A = (P^* E)(E P)$. Pick B to be EP and get the partial result.

Conversely, we may use the result of the previous exercise. If $y = (a_1, a_2, \ldots, a_n)$ is a vector in $\mathbb{F}^n$, then we have

$$y^* A y = \sum_{i,j} A_{ij} a_j \overline{a_i}$$

and

$$y^* A y = y^* B^* B y = (By)^* B y = \|By\|^2 \geq 0.$$

(d) Since T is self-adjoint, there's a basis β consisting of eigenvectors of T. For all $x \in \beta$, we have

$$U^2(x) = T^2(x) = \lambda^2 x.$$

If $\lambda = 0$, then we have $U^2(x) = 0$ and so $U(x) = 0 = T(x)$ since

$$\langle U(x), U(x) \rangle = \langle U^* U(x), x \rangle = \langle U^2(x), x \rangle = 0.$$

By the previous arguments we may assume that $\lambda > 0$. And this means that

$$0 = (U^2 - \lambda^2 I)(x) = (U + \lambda I)(U - \lambda I)(x).$$

But $\det(U + \lambda I)$ cannot be zero otherwise the negative value $-\lambda$ is an eigenvalue of U. So we have $U + \lambda I$ is invertible and $(U - \lambda I)(x) = 0$. Hence we get $U(x) = \lambda x = T(x)$. Finally since U and T meet on the basis β, we have $U = T$.

(e) We have T and U are diagonalizable since they are self-adjoint. Also, by the fact $TU = UT$ and Exercise 5.4.25, we may find a basis β consisting of eigenvectors of U and T. Say $x \in \beta$ is an eigenvector of T and U with respect to λ and μ, who are nonnegative since T and U are postive definite. Finally we get that all eigenvalue of TU is nonnegative since $TU(x) = \lambda \mu x$. So $TU = UT$ is also positive definite since they are self-adjoint by Exercise 6.4.4.

(f) Follow the notation of Exercise 6.4.17(b) and denote $y = (a_1, a_2, \ldots, a_n)$. We have

$$\langle T(\sum_{i=1}^{n} a_i e_i), \sum_{i=1}^{n} a_i e_i \rangle = \langle \sum_{i=1}^{n} (\sum_{j=1}^{n} A_{ij} a_j) e_i, \sum_{i=1}^{n} a_i e_i \rangle$$

$$\sum_{i,j} A_{ij} a_j \overline{a_i} = y^* A y = \langle L_A(y), y \rangle.$$

So the statement is true.

18. (a) We have T^*T and TT^* are self-adjoint. If λ is an eigenvalue with the eigenvector x, then we have $T^*T(x) = \lambda x$. Hence

$$\lambda = \langle T^*T(x), x \rangle = \langle T(x), T(x) \rangle \geq 0.$$

We get that T^*T is positive semidefinite by Exercise 6.4.17(a). By similar way we get the same result for TT^*.

(b) We prove that $N(T^*T) = N(T)$. If $x \in N(T^*T)$, we have

$$\langle T^*T(x), x \rangle = \langle T(x), T(x) \rangle = 0$$

and so $T(x) = 0$. If $x \in N(T)$, we have $T^*T(x) = T^*(0) = 0$.

Now we get that $\mathrm{null}(T^*T) = \mathrm{null}(T)$ and $\mathrm{null}(TT^*) = \mathrm{null}(T^*)$ since $T^{**} = T^*$. Also, we have $\mathrm{rank}(T) = \mathrm{rank}(T^*)$ by the fact

$$\mathrm{rank}([T]_\beta) = \mathrm{rank}([T]^*_\beta) = \mathrm{rank}([T^*]_\beta)$$

for some orthonormal basis β. Finally by Dimension Theorem we get the result

$$\mathrm{rank}(T^*T) = \mathrm{rank}(T) = \mathrm{rank}(T^*) = \mathrm{rank}(TT^*).$$

19. (a) It comes from that

$$\langle (T + U)(x), x \rangle = \langle T(x), x \rangle + \langle U(x), x \rangle > 0$$

and $(T + U)^* = T^* + U^* = T + U$.

(b) It comes from that

$$\langle (cT)(x), x \rangle = c \langle T(x), x \rangle > 0$$

and $(cT)^* = \bar{c}T^* = cT$.

(c) It comes from that

$$\langle T^{-1}(x), x \rangle = \langle y, T(y) \rangle > 0,$$

where $y = T^{-1}(x)$. Note that

$$(T^{-1})^*T^* = (TT^{-1})^* = I.$$

So we have $(T^{-1})^* = (T^*)^{-1} = T^{-1}$.

20. Check the condition one by one.

- $$\langle x + z, y \rangle' = \langle T(x + z), y \rangle$$
$$= \langle T(x), y \rangle + \langle T(z), y \rangle = \langle x, y \rangle' + \langle z, y \rangle'.$$

- $$\langle cx, y \rangle = \langle T(cx), y \rangle$$
$$= c \langle T(x), y \rangle = \langle x, y \rangle'.$$

- $$\overline{\langle x, y \rangle'} = \overline{\langle T(x), y \rangle}$$
$$= \langle y, T(x) \rangle = \langle T(y), x \rangle = \langle y, x \rangle'.$$

- $$\langle x, x \rangle' = \langle T(x), x \rangle > 0$$

if x is not zero.

187

21. As Hint, we check whether UT is self-adjoint with respect to the inner product $\langle x, y \rangle'$ or not. Denote F to be the operator UT with respect to the new inner product. Compute that

$$\langle x, F^*(y) \rangle' = \langle UT(x), y \rangle' = \langle TUT(x), y \rangle$$
$$= \langle T(x), UT(y) \rangle = \langle x, F(y) \rangle'$$

for all x and y. This means that UT is self-adjoint with respect to the new inner product. And so there's some orthonormal basis consisting of eigenvectors of UT and all the eigenvalue is real by the Lemma before Theorem 6.17. And these two properties is independent of the choice of the inner product. On the other hand, T^{-1} is positive definite by Exercie 6.4.19(c). So the function $\langle x, y \rangle'' := \langle T^{-1}(x), y \rangle$ is also a inner product by the previous exercise. Denote F' to be the operator TU with respect to this new inner product. Similarly, we have

$$\langle x, F'^*(y) \rangle'' = \langle TU(x), y \rangle'' = \langle U(x), y \rangle$$
$$= \langle T^{-1}(x), TU(y) \rangle = \langle x, F'(y) \rangle''$$

for all x and y. By the same argument we get the conclusion.

22. (a) For brevity, denote V_1 and V_2 to be the spaces with inner products $\langle \cdot, \cdot \rangle$ and $\langle \cdot, \cdot \rangle'$ respectively. Define $f_y(x) = \langle x, y \rangle'$ be a function from V_1 to $\mathbb{F}$. We have that $f_y(x)$ is linear for x on V_1. By Theorem 6.8 we have $f_y(x) = \langle T(x), y \rangle$ for some unique vector $T(x)$. To see T is linear, we may check that

$$\langle T(x+z), y \rangle = \langle x+z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$$
$$= \langle T(x), y \rangle + \langle T(z), y \rangle = \langle T(x) + T(z), y \rangle$$

and

$$\langle T(cx), y \rangle = \langle cx, y \rangle = c \langle x, y \rangle$$
$$c \langle T(x), y \rangle = \langle cT(x), y \rangle$$

for all x, y, and z.

(b) First, the operator T is self-adjoint since

$$\langle x, T^*(y) \rangle = \langle T(x), y \rangle = \langle x, y \rangle'$$
$$= \overline{\langle y, x \rangle} = \overline{\langle T(y), x \rangle} = \langle x, T(y) \rangle$$

for all x and y. Then T is positive definite on V_1 since

$$\langle T(x), x \rangle = \langle x, x \rangle' > 0$$

if x is not zero. Now we know that 0 cannot be an eigenvalue of T. So T is invertible. Thus T^{-1} is the unique operator such that

$$\langle x, y \rangle = \langle T^{-1}(x), y \rangle'.$$

By the same argument, we get that T^{-1} is positive definite on V_2. So T is also positive definite by Exercise 6.4.19(c).

23. As Hint, we denote V_1 and V_2 are the spaces with inner products $\langle \cdot, \cdot \rangle$ and $\langle \cdot, \cdot \rangle'$. By the definition of $\langle \cdot, \cdot \rangle'$, the basis β is orthonormal in V_2. So U is self-adjoint on V_2 since it has an orthonormal basis consisting of eigenvectors. Also, we get a special positive definite, and so self-adjoint, operator T_1 by Exercise 6.4.22 such that

$$\langle x, y \rangle' = \langle T(x), y \rangle.$$

We check that $U = T_1^{-1}U^*T_1$ by

$$\langle x, T_1^{-1}U^*T_1(y) \rangle = \langle T_1 U T_1^{-1}(x), y \rangle$$

$$= \langle U T_1^{-1}(x), y \rangle' = \langle T_1^{-1}(x), U(y) \rangle' = \langle x, U(y) \rangle$$

for all x and y. So we have $U = T_1^{-1}U^*T_1$ and so $T_1 U = U^*T_1$. Pick $T_2 = T_1^{-1}U^*$ and observe that it's self-adjoint. Pick $T_1' = T_1^{-1}$ to be a positive definite operator by Exercise 6.4.19(c). Pick $T_2' = U^*T_1$ to be a self-adjoint operator. Now we have $U = T_2 T_1 = T_1' T_2'$.

24. (a) Let
$$\beta = \{v_1, v_2, \ldots, v_n\}$$
and
$$\gamma = \{w_1, w_2, \ldots, w_n\}$$
be the two described basis. Denote A to be $[T]_\beta$. We have

$$T(w_1) = T\left(\frac{v_1}{\|v_1\|}\right) = \frac{A_{11}}{\|v_1\|}T(v_1) = A_{11}T(v_1).$$

Let t be the maximum integer such that $T(w_t)$ is an element in

$$\text{span}\{w_1, w_2, \ldots, w_t\}.$$

If $t = \dim(V)$, then we've done. If not, we have that

$$w_{t+1} = \frac{1}{L}\left(v_t - \sum_{j=1}^{t}\langle v_{t+1}, w_j \rangle w_j\right),$$

where

$$L = \left\|v_t - \sum_{j=1}^{t}\langle v_{t+1}, w_j \rangle w_j\right\|.$$

By the definition of w_i's we may define

$$W_i = \text{span}\{v_1, v_2, \ldots, v_n\} = \text{span}\{w_1, w_2, \ldots, w_2\}.$$

Now we have $T(w_t) \in W_t$ since

$$T(w_t) = \frac{1}{L}\left(T(v_t) - \sum_{j=1}^{t-1}\langle v_t, w_j \rangle T(w_j)\right)$$

and $T(v_t) \in W_t$ and $T(w_j) \in W_j \subset W_t$ for all $j < t$. This is a contradiction to our choice of i. So $[T]_\gamma$ is an upper triangular matrix.

(b) If the characteristic polynomial of T splits, we have an ordered basis β such that $[T]_\beta$ is upper triangular. Applying the previous argument, we get an orthonormal basis γ such that $[T]_\gamma$ is upper triangular.

6.5 Unitary and Orthogonal Operators and Their Matrices

1. (a) Yes. See Theorem 6.18.

   (b) No. Each rotation operator with nonzero angle is a counterexample.

   (c) No. A matrix is invertible if it's unitary. But an invertible matrix, $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ for example, may not be unitary.

   (d) Yes. It comes from the definition of unitarily equivalence.

   (e) No. For example, the idenetity matrix I is an unitary matrix but the sum $I + I$ is not unitary.

   (f) Yes. It's because that T is unitary if and only if $TT^* = T^*T = I$.

   (g) No. The basis β should be an orthonormal basis. For example, we have $T(a, b) = (b, a)$ is an orthogonal operator. But when we pick β to be
   $$\{(1,1), (1,0)\}$$
   we get that $[T]_\beta = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$ is not orthogonal.

   (h) No. Consider the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Its eigenvalues are 1. But it's not orthogonal.

   (i) No. See Theorem 6.18.

2. Just follow the process of diagonalization. But remember that if the dimension of some eigenspace is more than 1, we should choose an orthonormal basis on it.

   (a)
   $$P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, D = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}.$$

   (b)
   $$P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}, D = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}.$$

   (c)
   $$P = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & \sqrt{2} \\ i+1 & -\frac{i+1}{\sqrt{2}} \end{pmatrix}, D = \begin{pmatrix} 8 & 0 \\ 0 & -1 \end{pmatrix}.$$

(d)

$$P = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{\sqrt{2}}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix}, D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

(e)

$$P = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{\sqrt{2}}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix}, D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

3. If T and U are unitary [orthogonal] operators, then we have

$$\|TU(x)\| = \|U(x)\| = \|x\|.$$

4. Pick the standard basis β and compute the matrix representation $[T_x]_\beta = (z)$. This means that $T_z^* = T_{\bar{z}}$. So it always would be normal. However, it would be self-adjoint only when z is real. And it would be unitary only when $|z| = 1$.

5. For these problem, try to diagonalize the matrix which is not diagonalized yes. And check whether it can be diagonalized by an orthonormal basis.

   (a) No. They have different eigenvalues.
   (b) No. Their determinant is different.
   (c) No. They have different eigenvalues.
   (d) Yes. We have

   $$\begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{pmatrix}^* \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix}.$$

   (e) No. One is symmetric but the other is not.

6. If T is unitary, we must have

$$0 = \|T(f)\|^2 - \|f\|^2 = \int_0^1 |h|^2 |f|^2 dt - \int_0^1 |f|^2 dt$$

$$= \int_0^1 (1 - |h|^2) |f|^2 dt$$

for all $f \in V$. Pick $f = (1 - |h|^2)^{\frac{1}{2}}$ and get $1 - |h|^2 = 0$ and so $|h| = 1$. Conversely, if $|h| = 1$, we have

$$\|T(f)\|^2 - \|f\|^2 = \int_0^1 |h|^2 |f|^2 dt - \int_0^1 |f|^2 dt$$

$$= \int_0^1 (1 - |h|^2) |f|^2 dt = 0$$

and so T is unitary.

7. By the Corollary 2 after Theorem 6.18, we may find an orthonormal basis β such that

$$[T]_\beta = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Also, since the eigenvalue λ_i has its absolute value 1, we may find some number μ_i such that $\mu_i^2 = \lambda_i$ and $|\mu_i| = 1$. Denote

$$D = \begin{pmatrix} \mu_1 & 0 & \cdots & 0 \\ 0 & \mu_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \mu_n \end{pmatrix}$$

to be an unitary operator. Now pick U to be the matrix whose matrix representation with respect to β is D. Thus U is unitary and $U^2 = T$.

8. Exercise 6.4.10 says that $((T - iI)^{-1})* = T + iI$. So check that

$$[(T + iI)(T - iI)^{-1}]^*(T + iI)(T - iI)^{-1}$$

$$= ((T - iI)^{-1})^*(T + iI)^*(T + iI)(T - iI)^{-1}$$

$$= (T + iI)^{-1}(T - iI)(T + iI)(T - iI)^{-1}$$

$$= (T + iI)^{-1}(T + iI)(T - iI)(T - iI)^{-1} = I.$$

Use Exercise 2.4.10 we get that the operator is unitary.

9. The operator U may not be unitary. For example, let $U(a, b) = (a + b, 0)$ be an operator on $\mathbb{C}^2$. Pick the basis $\{(1, 0), (0, 1)\}$ and we may observe that

$$\|U(1, 0)\| = \|U(0, 1)\| = 1 = \|(1, 0)\| = \|(0, 1)\|.$$

But it is not unitary since

$$\|U(1, 1)\| = 1 \neq \|(1, 1)\| = \sqrt{2}.$$

10. Exercise 2.5.10 says that $\operatorname{tr}(A) = \operatorname{tr}(B)$ is A is similar to B. And we know that A may be diagonalized as $P^*AP = D$ by Theorem 6.19 and Theorem 6.20. Here D is a diagonal matrix whose diagonal entries consist of all eigenvalues. This means

$$\operatorname{tr}(A) = \operatorname{tr}(D) = \sum_{i=1}^{n} \lambda_i$$

and

$$\operatorname{tr}(A^*A) = \operatorname{tr}((PDP^*)^*(PDP^*))$$

$$= \operatorname{tr}(PD^*DP^*) = \operatorname{tr}(D^*D) = \sum_{i=1}^{n} |\lambda_i|^2.$$

11. Extend $\{(\frac{1}{3}, \frac{2}{3}, \frac{2}{3})\}$ to be a basis and do the Gram-Schmidt process to get an orthonormal basis. The extended basis could be

$$\{(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}), (0, 1, 0), (0, 0, 1)\}$$

and the othonormal basis would be

$$\{(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}), (-\frac{2}{3\sqrt{5}}, \frac{5}{3\sqrt{5}}, -\frac{2^2}{3\sqrt{5}}), (-\frac{2}{\sqrt{5}}, 0, \frac{1}{\sqrt{5}})\}.$$

So the matrix could be

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

12. By Theorem 6.19 and Theorem 6.20 we know that A may be diagonalized as $P^*AP = D$. Here D is a diagonal matrix whose diagonal entries consist of all eigenvalues. Now we have

$$\det(A) = \det(PDP^*) = \det(D) = \prod_{i=1}^{n} \lambda_i.$$

13. The necessity is false. For example, the two matrices $\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$ and

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

are similar. But they are not unitary since one is symmetric but the other is not.

14. We may write $A = P^*BP$. Compute

$$\langle L_A(x), x \rangle = \langle L_{P^*BP}(x), x \rangle$$

$$= \langle L_P^* L_B L_P(x), x \rangle = \langle L_B(L_P(x)), L_P(x) \rangle.$$

If A is positive definite, for each vector y we may find some x such that $L_P(x) = y$ since P is invertible. Also, we have

$$\langle L_B(y), y \rangle = \langle L_B(L_P(x)), L_P(x) \rangle = \langle L_A(x), x \rangle > 0.$$

If B is positive definite, we may check that

$$\langle L_A(x), x \rangle \langle L_B(L_P(x)), L_P(x) \rangle > 0.$$

15. (a) We have
$$\|U_W(x)\| = \|U(x)\| = \|x\|$$

and so U_W is an unitary operator on W. Also, the equality above implies that U_W is injection. Since W is finite-dimensional, we get that U_W is surjective and $U(W) = W$.

(b) For each elment $w \in W$ we have $U(y) = w$ for some $y \in W$ by the previous argument. Now let x be an element in $W^\perp$. We have $U(x) = w_1 + w_2$ for some $w_1 \in W$ and $w_2 \in W^\perp$ by Exercise 6.2.6. Since U is unitary, we have some equalities
$$\|y\|^2 = \|w\|^2,$$
$$\|x\|^2 = \|w_1 + w_2\|^2 = \|w_1\|^2 + \|w_2\|^2$$

by Exercise 6.1.10. However, we also have that $U(x+y) = 2w_1 + w_2$. So we have that
$$0 = \|x + y\|^2 - \|2w_1 + w_2\|^2$$
$$= \|x\|^2 + \|y\|^2 - 4\|w_1\|^2 - \|w_2\|^2 = -2\|w_1\|^2.$$

This means that $w_1 = 0$ and so $U(x) = w_2 \in W^\perp$.

16. This example show the finiteness in the previous exercise is important. Let V be the space of sequence defined in Exercise 6.2.23. Also use the notation e_i in the same exercise. Now we define a unitary operator U by

$$\begin{cases} U(e_{2i+1}) = e_{2i-1} & \text{if} \quad i > \quad ; \\ U(e_1) = e_2 & \quad ; \\ U(e_{2i}) = U(e_{2i+2}) & \text{if} \quad i > 0 \quad . \end{cases}$$

It can be check that $\|U(x)\| = \|x\|$ and U is surjective. So U is an unitary operator. We denote W to be the subspace

$$\text{span}\{e_2, e_4, e_6, \ldots\}$$

and so we have
$$W^\perp = \{e_1, e_3, e_5, \ldots\}.$$

Now, W is a U-invariant subspace by definition. However, we have $e_2 \notin U(W)$ and $W^\perp$ is not U-invariant since $U(e_1) = e_2 \notin W^\perp$.

17. Let A be an unitary and upper triangular matrix. For arbitrary indices i and j such that $i > j$. We have $A_{ij} = 0$ since A is upper triangular. But we also have that $A_{ji} = \overline{A_{ij}} = 0$. So A is a diagonal matrix.

18. Write $A \sim B$ to say that A is unitarily equivalent to B. Check the three conditions in the Appendix A.

reflexivity Since $A = I^* A I$, we get $A \sim A$.

symmetry If $A \sim B$, we have $B = P^*AP$ and so $A = PBP^* = (P^*)^*BP^*$. This means that $B \sim A$.

transitivity If $A \sim B$ and $B \sim C$, we have $B = P^*AP$ and $C = Q^*BQ$. This means that

$$C = Q^*BQ = Q^*P^*APQ = (PQ)^*A(PQ)$$

and so $A \sim C$.

19. By Exercise 6.1.10 we have

$$\|U(v_1 + v_2)\|^2 = \|v_1 - v_2\|^2$$

$$= \|v_1\|^2 + \|v_2\|^2 = \|v_1 + v_2\|^2.$$

20. (a) If it's a complex inner product space, we have

$$\langle U(x), U(y) \rangle = \sum_{k=1}^{4} i^k \|U(x) + iU(y)\|^2$$

$$\sum_{k=1}^{4} i^k \|U(x + iy)\|^2 = \sum_{k=1}^{4} i^k \|x + iy\|^2 = \langle x, y \rangle.$$

If it's a real inner product space, we may also use the above equality but take the summation only over $k = 2, 4$.

(b) Since $U(x + y) = U(x)$ if $x \in W$ an $y \in W^{\perp}$, we know that $R(U) = U(W)$. We get the desired result by applying the previous argument.

(c) Just extend the set
$$\{v_1, v_2, \ldots, v_k\}$$

to be an orthonormal basis

$$\gamma = \{v_1, v_2, \ldots, v_n\}$$

for V, where $n = \dim(V)$. First we know that $U(v_j) = 0$ if $j > k$. So the j-th column of $[U]_\gamma$ is zero. On the other hand, if we write $A = [U]_\gamma$ we have
$$U(v_j) = \sum_{i=1}^{n} U_{ij}v_i.$$

So the first k columns is orthonormal since

$$0 = \langle U(v_s), U(v_t) \rangle = \langle \sum_{i=1}^{n} U_{is}v_i, \sum_{i=1}^{n} U_{it}v_i \rangle$$

$$= \sum_{i=1}^{n} U_{is}\overline{U_{it}} = (Ae_t)^* Ae_s$$

195

and

$$1 = \langle U(v_s), U(v_s) \rangle = \langle \sum_{i=1}^{n} U_{is} v_i, \sum_{i=1}^{n} U_{is} v_i \rangle$$

$$= \sum_{i=1}^{n} U_{is} \overline{U_{is}} = (Ae_s)^* Ae_s.$$

(d) Since V is finite-dimensional inner product space, we have $R(U)^\perp \oplus R(U) = V$. And so by Exercise 6.5.20(b) we get the desired result.

(e) First, T is well-defined since the set β defined in the previous question is a basis. To show that $T = U^*$, it's sufficient to check that

$$\langle U(x), y \rangle = \langle x, T(y) \rangle$$

for all x and y in β by Exercise 6.1.9. We partition β into two parts X and Y, who consist of all $U(v_i)$'s and w_i's.

- If $x = U(v_i), y = U(v_j) \in X$, we have

$$\langle U(v_i), T(U(v_j)) \rangle = \langle U(v_i), v_j \rangle = \langle U^2(v_i), U(v_j) \rangle$$

by Exercise 6.5.20(a).

- If $x = U(v_i) \in X$ and $y = w_j \in Y$, we have

$$\langle U(v_i), T(w_j) \rangle = \langle U(v_i), 0 \rangle$$

$$= 0 = \langle U^2(v_i), w_j \rangle.$$

- If $x = w_i \in X$ and $y = U(v_j) \in Y$, we have

$$\langle w_i, T(U(v_j)) \rangle = \langle w_i, v_j \rangle = U(w_i), U(v_j) \rangle.$$

- If $x = w_i, y = w_j \in Y$, we have

$$\langle w_i, T(w_j) \rangle = \langle w_i, 0 \rangle = 0 = \langle U(w_i), w_j \rangle.$$

(f) Take the subspace W' to be $R(U)$. Thus we have $T((W')^\perp) = \{0\}$ by the definition of T. Also, we may write an element x in $R(U)$ to be

$$x = \sum_{i=1}^{k} a_i U(v_i).$$

Since the set of $U(v_i)$'s is orthonormal, we have

$$\|x\|^2 = \sum_{i=1}^{k} |a_i|^2 = \|\sum_{i=1}^{k} a_i v_i\|^2 = \|T(x)\|^2.$$

21. Since A is unitarily equivalent to B, we write $B = P^* A P$ for some unitary matrix P.

(a) Compute

$$\mathrm{tr}(B^* B) = \mathrm{tr}((P^* AP)^*(P^* AP)) = \mathrm{tr}(P^* A^* AP) = \mathrm{tr}(A^* A).$$

(b) We compute the trace of $A^* A$ and get

$$\mathrm{tr}(A^* A) = \sum_{i=1}^{n} (A^* A)_{ii}$$

$$= \sum_{i=1}^{n} \sum_{k=1}^{n} (A^*)_{ik} A_{ki} = \sum_{i,j} |A_{ij}|^2.$$

Use the result in the previous argument we get the conclusion.

(c) By the previous argument, they are not unitarily equivalent since

$$|1|^2 + |2|^2 + |2|^2 + |i|^2 = 10$$

is not equal to

$$|i|^2 + |4|^2 + |1|^2 + |1|^2 = 19.$$

22. (a) Let $f(x) = x + t$ be a translation. We may check it's a rigid motion by

$$\|f(x) - f(y)\| = \|(x + t) - (y + t)\| = \|x - y\|.$$

(b) Let f, g be two rigid motion, we have

$$\|fg(x) - fg(y)\| = \|g(x) - g(y)\| = \|x - y\|.$$

So the composition of f and g is again a rigid motion.

23. We define T to be $T(x) = f(x) - f(0)$. By the proof of Theorem 6.22, we know that T is an unitary operator. Also, by Theorem 6.22 we know f is surjective since it's composition of two invertible functions. Hence we may find some element t such that $f(t) = 2f(0)$. Now let $g(x) = x + t$. Since T is linear, we have

$$T \circ g(x) = T(x + t) = T(x) + T(t)$$

$$= f(x) - f(0) + f(t) - f(0) = f(x).$$

Finally, if $f(x) = T(x + t) = U(x + v_0)$ for some unitary operator U and some element v_0. We'll have

$$T(-v_0 + t) = U(-v_0 + v_0) = 0.$$

Since T is unitary and hence injective, we know that $t = v_0$. And thus U must equal to T. So this composition is unique.

24. (a) First, the composition of two unitary operators is again an unitary operator. So UT is an unitary operator. Since $\det(U) = \det(T) = -1$, we have

$$\det(UT) = \det(U)\det(T) = 1.$$

This means that UT must be a rotation by Theorem 6.23.

(b) It's similar to the previous argumemnt. And now we have

$$\det(UT) = \det(TU) = \det(T)\det(U) = 1\cdot(-1) = -1.$$

So they are reflections.

25. By the proof of Theorem 6.23 we know that the matrix representations of T and U with repect to the standard basis α are

$$[T]_\alpha = \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix}$$

and

$$[U]_\alpha = \begin{pmatrix} \cos 2\psi & \sin 2\psi \\ \sin 2\psi & -\cos 2\psi \end{pmatrix}.$$

So we have

$$[UT]_\alpha = [U]_\alpha[T]_\alpha = \begin{pmatrix} \cos 2(\psi-\phi) & -\sin 2(\psi-\phi) \\ \sin 2(\psi-\phi) & \cos 2(\psi-\phi) \end{pmatrix}.$$

Hence UT is a rotation by the angle $2(\psi-\phi)$.

26. Here we have

$$[T]_\alpha = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

and

$$[U]_\alpha = \begin{pmatrix} \cos 2\psi & \sin 2\psi \\ \sin 2\psi & -\cos 2\psi \end{pmatrix}.$$

(a) Compute

$$[UT]_\alpha = [U]_\alpha[T]_\alpha = \begin{pmatrix} \cos 2(\psi-\frac{\phi}{2}) & \sin 2(\psi-\frac{\phi}{2}) \\ \sin 2(\psi-\frac{\phi}{2}) & -\cos 2(\psi-\frac{\phi}{2}) \end{pmatrix}.$$

So the angle is $\psi - \frac{\phi}{2}$.

(b) Compute

$$[TU]_\alpha = [T]_\alpha[U]_\alpha = \begin{pmatrix} \cos 2(\psi+\frac{\phi}{2}) & \sin 2(\psi+\frac{\phi}{2}) \\ \sin 2(\psi+\frac{\phi}{2}) & -\cos 2(\psi+\frac{\phi}{2}) \end{pmatrix}.$$

So the angle is $\psi + \frac{\phi}{2}$.

27. (a) We may write
$$(x \quad y) \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Diagonalize the matrix and get
$$(x \quad y) P^* \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} P \begin{pmatrix} x \\ y \end{pmatrix},$$

where
$$P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

So we have
$$\begin{pmatrix} x' \\ y' \end{pmatrix} = P \begin{pmatrix} x \\ y \end{pmatrix}.$$

(b) Diagonalize $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ and get
$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

(c) Diagonalize $\begin{pmatrix} 1 & -6 \\ -6 & -4 \end{pmatrix}$ and get
$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \frac{1}{\sqrt{13}} \begin{pmatrix} 2 & 3 \\ 3 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

(d) Diagonalize $\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ and get
$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

(e) Diagonalize $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ and get
$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

28. Denote $(X')^t = (x'y', z')$. Then we have
$$X' = PX,$$

where P is the matrix in the solution of Exercise 6.5.2(e).

29. (a) We have the formula

$$\|v_k\| u_k = v_k = w_k - \sum_{j=1}^{k} \frac{\langle w_k, v_j \rangle}{\|v_j\|^2} v_j$$

$$= w_k - \sum_{j=1}^{k} \frac{\langle w_k, v_j \rangle}{\|v_j\|} u_j = w_k - \sum_{j=1}^{k} \langle w_k, u_j \rangle u_j.$$

(b) It directly comes from the formula above and some computation.

(c) We have

$$w_1 = (1,1,0), w_2 = (2,0,1), w_3 = (2,2,1)$$

and

$$v_1 = (1,1,0), v_2 = (1,-1,1), v_3 = \left(-\frac{1}{3}, \frac{1}{3}, \frac{2}{3}\right)$$

by doing the Gram-Schmidt process. This means that we have

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

Then we may also compute

$$\|v_1\| = \sqrt{2}, \|v_2\| = \sqrt{3}, \|v_3\| = \frac{\sqrt{2}}{\sqrt{3}}.$$

Now we have

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

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

$$= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} \sqrt{2} & \sqrt{2} & 2^{\frac{3}{2}} \\ 0 & \sqrt{3} & \frac{1}{\sqrt{3}} \\ 0 & 0 & \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix}.$$

Here the we have

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix}$$

and

$$R = \begin{pmatrix} \sqrt{2} & \sqrt{2} & 2^{\frac{3}{2}} \\ 0 & \sqrt{3} & \frac{1}{\sqrt{3}} \\ 0 & 0 & \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix}.$$

(d) First that Q_1, Q_2 and R_1, R_2 are invertible otherwise A cannot be invertible. Also, since Q_1, Q_2 is unitary, we have $Q_1^* = Q_1^{-1}$ and $Q_2^* = Q_2^{-1}$. Now we may observe that $Q_1 Q_2^* = R_2 R_1^{-1}$ is an unitary matrix. But $R_2 R_1^{-1}$ is upper triangular since R_2 and the inverse of an upper triangular matrix R_1 are triangular matrices. So $D = R_2 R_1^{-1}$ is both upper triangular and unitary. It could only be a unitary diagonal matrix.

(e) Denote b by $\begin{pmatrix} 1 \\ 11 \\ -1 \end{pmatrix}$. Now we have $A = QR = b$. Since Q is unitary, we have $R = Q^* b$. Now we have

$$
\begin{pmatrix} \sqrt{2} & \sqrt{2} & 2^{\frac{3}{2}} \\ 0 & \sqrt{3} & \frac{1}{\sqrt{3}} \\ 0 & 0 & \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix} = R = Q^* b = \begin{pmatrix} 3 \cdot 2^{\frac{3}{2}} \\ -\frac{11}{\sqrt{3}} \\ \frac{2^{\frac{5}{2}}}{\sqrt{3}} \end{pmatrix}.
$$

Then we may solve it to get the answer $x = 3$, $y = -5$, and $z = 4$.

30. We may write
$$
\beta = \{v_1, v_2, \ldots, v_n\}
$$
and
$$
\gamma = \{u_1, u_2, \ldots, u_n\}.
$$
We have that $Q = [I]_\gamma^\beta$. Now if β is orthonormal, we may compute

$$
= u_j = \sum_{i=1}^n Q_{ij} v_i.
$$

Thus we know that the inner product of u_s and u_t would be

$$
\langle u_s, u_t \rangle = \langle \sum_{i=1}^n Q_{is} v_i, \sum_{i=1}^n Q_{it} v_i \rangle
$$

$$
= \sum_{i=1}^n Q_{is} \overline{Q_{it}},
$$

the value of inner product of the s-th and the t-th columns of Q. So it would be 1 if $s = t$ and it would be 0 if $s \neq t$. Finally the converse is also true since $Q^* = [I]_\beta^\gamma$ is also an unitary matrix.

31. (a) Check that
$$
H_u(x + cy) = x + cy - 2\langle x + cy, u \rangle u
$$
$$
= (x - 2\langle x, u \rangle u) + c(y - 2\langle y, u \rangle u)
$$
$$
= H_u(x) + cH_u(y).
$$

(b) Compute
$$H_u(x) - x = -2\langle x, u \rangle u.$$
The value would be zero if and only if x is orthogonal to u since u is not zero.

(c) Compute
$$H_u(u) = u - 2\langle u, u \rangle u = u - 2u = -u.$$

(d) We check $H_u^* = H_u$ by computing
$$\langle x, H_u^*(y) \rangle = \langle H_u(x), y \rangle = \langle x - 2\langle x, u \rangle u \rangle$$
$$= \langle x, y \rangle - 2\langle x, u \rangle \cdot \langle y, u \rangle$$
and
$$\langle x, H_u(y) \rangle = \langle x, y - 2\langle y, u \rangle u \rangle$$
$$= \langle x, y \rangle - 2\langle x, u \rangle \cdot \langle y, u \rangle.$$
Also, compute
$$H_u^2(x) = H_u(x - 2\langle x, u \rangle u)$$
$$H_u(x) - 2\langle x, u \rangle H_u(u)$$
$$(x - 2\langle x, u \rangle u) + 2\langle x, u \rangle u = x.$$
Combining $H_u = H_u^*$ and $H_u^2 = I$, we have $H_u H_u^* = H_u^* H_u = I$ and so H_u is unitary.

32. (a) We pick θ to be the value $\langle x, y \rangle$. Thus we have
$$\overline{\langle x, \theta y \rangle} = \overline{\langle x, \langle x, y \rangle y \rangle}$$
$$= \overline{\langle y, x \rangle \langle x, y \rangle} = \langle x, y \rangle \langle y, x \rangle$$
$$= \langle x, \langle x, y \rangle y \rangle = \langle x, \theta y \rangle$$
and so the value is real. Now pick u to be the normalized vector $\frac{x - \theta y}{\|x - \theta y\|}$. And compute
$$H_u(x) = x - 2\langle x, u \rangle u = x - 2\langle x, \frac{x - \theta y}{\|x - \theta y\|} \rangle \frac{x - \theta y}{\|x - \theta y\|}$$
$$= x - \frac{2}{\|x - \theta y\|^2}\langle x, x - \theta y \rangle(x - \theta y)$$
$$= x - \frac{2\|x\|^2 - 2\langle x, \theta y \rangle}{\|x - \theta y\|^2}(x - \theta y)$$
$$= x - \frac{\|x\|^2 - 2\langle x, y \rangle + \|\theta y\|^2}{\|x - \theta y\|^2}(x - \theta y)$$
$$= x - x + \theta y = \theta y.$$

202

(b) Pick u to be the normalized vector $\frac{x-y}{\|x-y\|}$. Compute

$$H_u(x) = x - 2\langle x, u \rangle u = x - 2\langle x, \frac{x-y}{\|x-y\|}\rangle \frac{x-y}{\|x-y\|}$$

$$= x - \frac{2}{\|x-y\|^2}\langle x, x-y\rangle(x-y)$$

$$= x - \frac{2\|x\|^2 - 2\langle x, y\rangle}{\|x-y\|^2}(x-y)$$

$$= x - \frac{\|x\|^2 - 2\langle x, y\rangle + \|y\|^2}{\|x-y\|^2}(x-y)$$

$$= x - x + y = y.$$

6.6 Orthogonal Projections and the Spectral Theorem

1. (a) No. Orthogonal projection is self-adjoint by Theorem 6.24. But for general projection the statement is not true. For example, the transformation $T(a, b) = (a + b, 0)$ is a projection which is not self-adjoint.

   (b) Yes. See the paragraph after Definition of "orthogonal projection".

   (c) Yes. This is the result of the Spectral Theorem.

   (d) No. It's true for orthogonal projection but false for general projection. For example, the the transformation $T(a, b) = (a + b, 0)$ is a projection on W. But we have $T(0, 1) = (1, 0)$ is not the point closest to $(0, 1)$ since $(0, 0)$ is much closer.

   (e) No. An unitary operator is usually invertible. But an projection is generally not invertible. For example, the mapping $T(a, b) = (a, 0)$.

2. We could calculate the projection of $(1, 0)$ and $(0, 1)$ are

$$\frac{\langle (1, 0), (1, 2) \rangle}{\|(1, 2)\|^2}(1, 2) = \frac{1}{5}(1, 2)$$

and

$$\frac{\langle (0, 1), (1, 2) \rangle}{\|(1, 2)\|^2}(1, 2) = \frac{2}{5}(1, 2)$$

by Theorem 6.6. So we have

$$[T]_\beta = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}.$$

On the other hand, we may do the same on $(1,0,0)$, $(0,1,0)$, and $(1,0,0)$ with respect to the new subspace $W = \text{span}(\{(1,0,1)\})$. First compute

$$\frac{\langle (1,0,0),(1,0,1)\rangle}{\|(1,0,1)\|^2}(1,0,1) = \frac{1}{2}(1,0,1),$$

$$\frac{\langle (0,1,0),(1,0,1)\rangle}{\|(1,0,1)\|^2}(1,0,1) = 0(1,0,1),$$

and

$$\frac{\langle (0,0,1),(1,0,1)\rangle}{\|(1,0,1)\|^2}(1,0,1) = \frac{1}{2}(1,0,1).$$

Hence the matrix would be

$$[T]_\beta = \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

3. The first and the third step comes from the Spectral theorem and the fact that these matrices are self-adjoint or at least normal. So we only do the first two steps. Also, we denote the matrix E_{ij} to be a matrix, with suitable size, whose ij-entry is 1 and all other entries are zero. Finally, it's remarkble that that the matrices P and D are different from the each questions. They are defined in Exercise 6.5.2.

   (a) Let $A_3 = P^* E_{11} P$ and $A_{-1} = P^* E_{22} P$. Then we have $T_3 = L_{A_3}$, $T_{-1} = L_{A_{-1}}$ and
   $$L_A = 3T_3 - 1T_{-1}.$$

   (b) Let $A_i = P^* E_{11} P$ and $A_i = P^* E_{22} P$. Then we have $T_{-i} = L_{A_{-i}}$, $T_i = L_{A_i}$ and
   $$L_A = -iT_{-i} + iT_i.$$

   (c) Let $A_8 = P^* E_{11} P$ and $A_{-1} = P^* E_{22} P$. Then we have $T_8 = L_{A_8}$, $T_{-1} = L_{A_{-1}}$ and
   $$L_A = 8T_8 - 1T_{-1}.$$

   (d) Let $A_4 = P^* E_{11} P$ and $A_{-2} = P^* (E_{22} + E_{33}) P$. Then we have $T_4 = L_{A_4}$, $T_{-2} = L_{A_{-2}}$ and
   $$L_A = 4T_4 - 2T_{-2}.$$

   (e) Let $A_4 = P^* E_{11} P$ and $A_1 = P^* (E_{22} + E_{33}) P$. Then we have $T_4 = L_{A_4}$, $T_1 = L_{A_1}$ and
   $$L_A = 4T_4 + 1T_1.$$

4. Since T is an orthogonal projection, we have $N(T) = R(T)^\perp$ and $R(T) = N(T)^\perp$. Now we want to say that $N(I - T) = R(T) = W$ and $R(I -$

$T) = N(T) = W^\perp$ and so $I - T$ is the orthogonal projection on $W^\perp$. If $x \in N(I - T)$, we have $x = T(x) \in R(T)$. If $T(x) \in R(T)$, we have

$$(I - T)T(x) = T(x) - T^2(x) = T(x) - T(x) = 0.$$

So we have the first equality. Next, if $(I - T)(x) \in R(I - T)$ we have

$$T(I - T)(x) = T(x) - T^2(x) = T(x) - T(x) = 0.$$

If $x \in N(T)$ we have $T(x) = 0$ and so $x = (I - T)(x) \in R(I - T)$. So the second equality also holds.

5. (a) Since T is an orthogonal projection, we may write $V = R(T) \oplus R(T)^\perp$. So for each $x \in V$ we could write $x = u + v$ such that $u \in R(T)$ and $v \in R(T)^\perp$. So we have

$$\|T(u + v)\| = \|u\| \le \|u + v\| = \|x\|.$$

The example for which the in which the inequality does not hold is $T(a, b) = (a + b, 0)$, since we have

$$\|T(1, 1)\| = \|(2, 0)\| = 2 > \|(1, 1)\| = \sqrt{2}.$$

Finally, if the equality holds for all $x \in V$, then we have $\|u\| = \|u + v\|$. Since u and v are orthogonal, we have

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

So the equality holds only when $v = 0$. This means that x is always an element in $R(T)$ and so $R(T) = V$. More precisely, T is the idenetity mapping on V.

(b) If T is a projection on W along W', we have $V = W \oplus W'$. So every vector $x \in V$ could be written as $x = u + v$ such that $u \in W$ and $v \in W'$. If $W' \ne W^\perp$, we may find some $u \in W$ and $v \in W'$ such that they are not orthogonal. So $\langle u, v \rangle$ is not zero. We may pick $t = \frac{2\|v\|^2}{2\mathrm{Re}\langle u, v \rangle}$ and calculate that
$$\|T(tu + v)\|^2 = \|tu\|^2.$$

But now we have

$$\|tu + v\|^2 = \|tu\|^2 + 2\mathrm{Re}\langle tu, v \rangle + \|v\|^2$$

$$= \|tu\|^2 - \|v\|^2 < \|T(tu + v)\|^2.$$

So T must be an orthogonal projection.

6. It's enough to show that $R(T)^\perp = N(T)$. If $x \in R(T)^\perp$, we have

$$\langle T(x), T(x) \rangle = \langle x, T^*T(x) \rangle = \langle x, T(T^*(x)) \rangle = 0$$

and so $T(x) = 0$. If $x \in N(T)$, we have

$$\langle x, T(y) \rangle = \langle T^*(x), y \rangle = 0$$

since $T^*(x) = 0$ by Theorem 6.15(c). Hence now we know that $R(T)^\perp = N(T)$ and T is an orthogonal projection.

7. (a) It comes from Theorem 6.25(c).

   (b) If

   $$T_0 = T^n = \sum_{i=1}^{k} \lambda_i^n T_i.$$

   Now pick arbitrary eigenvector v_i for the eigenvalue λ_i. Then we have

   $$0 = T(v_i) = \left(\sum_{i=1}^{k} \lambda_i^n T_i \right)(v_i) = \lambda_i^n v_i.$$

   This means that $\lambda_i^n = 0$ and so $\lambda_i = 0$ for all i. Hence we know that

   $$T = \sum_{i=1}^{k} \lambda_i T_i = T_0.$$

   (c) By the Corollary 4 after the Spectral Theorem, we know that $T_i = g_i(T)$ for some polynomial g_i. This means that U commutes with each T_i if U commutes with T. Conversely, if U commutes with each T_i, we have

   $$TU = \left(\sum_{i=1}^{k} \lambda_i T_i \right) U = \sum_{i=1}^{k} \lambda_i T_i U$$

   $$= \sum_{i=1}^{k} \lambda_i U T_i = U \left(\sum_{i=1}^{k} \lambda_i T_i \right) = UT.$$

   (d) Pick

   $$U = \sum_{i=1}^{k} \lambda_i^{\frac{1}{2}} T_i,$$

   where $\lambda_i^{\frac{1}{2}}$ is an arbitrary square root of λ_i.

   (e) Since T is a mapping from V to V, T is invertible if and only if $N(T) = \{0\}$. And $N(T) = \{0\}$ is equivalent to that 0 is not an eigenvalue of T.

   (f) If every eigen value of T is 1 or 0. Then we have $T = 0T_0 + 1T_1 = T_1$, which is a projectoin. Conversely, if T is a projection on W along W', we may write any element in V as $u + v$ such that $u \in W$ and $v \in W'$. And if λ is an eigenvalue, we have

   $$u = T(u + v) = \lambda(u + v)$$

   and so

   $$(1 - \lambda)u = \lambda v.$$

   Then we know that the eigenvalue could only be 1 or 0.

(g) It comes from the fact that

$$T^* = \sum_{i=1}^{k} \overline{\lambda_i T_i}.$$

8. It directly comes from that $T^* = g(T)$ for some polynomial g.

9. (a) Since U is an operator on a finite-dimensional space, it's sufficient to we prove that $R(U^*U)^\perp = N(U^*U)$ and U^*U is a projection. If $x \in R(U^*U)^\perp$, we have

$$\langle x, U^*U(x) \rangle = \langle U(x), U(x) \rangle = 0$$

and so $U^*U(x) = U^*(0) = 0$. If $x \in N(U^*U)$, we have

$$\langle x, U^*U(y) \rangle = U^*U(x), y \rangle = 0$$

for all y. This means that $x \in R(U^*U)^\perp$. Now we know that $V = R(U^*U) \oplus N(U^*U)$ and we can write element in V as $p + q$ such that $p \in R(U^*U)$ and $N(U^*U)$. Check that

$$U^*U(p + q) = U^*U(p) = p$$

by the definition of U^* in Exercise 6.5.30(e). Hence it's an orthogonal projection.

(b) Use the notation in Exercise 6.5.20. Let

$$\alpha = \{v_1, v_2, \ldots, v_k\}$$

be an orthonormal basis for W. Extend it to be an orthonormal basis

$$\gamma = \{v_1, v_2, \ldots, v_n\}$$

for V. Now check that

$$UU^*U(v_i) = 0 = U(v_i)$$

if $i > k$ and
$$UU^*U(v_i) = UU^*(U(v_i)) = U(v_i)$$

if $i \le k$ by the definition of U^* in Exercise 6.5.20(e). They meet on a basis and so they are the same.

10. We use induction on the dimension n of V. If $n = 1$, U and T will be diagonalized simultaneously by any orthonormal basis. Suppose the statement is true for $n \le k - 1$. Consider the case $n = k$. Now pick one arbitrary eigenspace $W = E_\lambda$ of T for some eigenvalue λ. Note that W is T-invariant naturally and U-invariant since

$$TU(w) = UT(w) = \lambda U(w)$$

for all $w \in W$. If $W = V$, then we may apply Theorem 6.16 to the operator U and get an orthonormal basis β consisting of eigenvectors of U. Those vectors will also be eigenvectors of T. If W is a proper subspace of V, we may apply the induction hypothesis to T_W and U_W, which are normal by Exercise 6.4.7 and Exercise 6.4.8, and get an orthonormal basis β_1 for W consisting of eigenvectors of T_W and U_W. So those vectors are also eigenvectors of T and U. On the other hand, we know that $W^\perp$ is also T- and U-invariant by Exercise 6.4.7. They are also normal operators by Exercise 6.4.7(d). Again, by applying the induction hypothesis we get an orthonormal basis β_2 for $W^\perp$ consisting of eigenvectors of T and U. Since V is finite dimentional, we know that $\beta = \beta_1 \cup \beta_2$ is an orthonormal basis for V consisting of eigenvectors of T and U.

11. By Theorem 6.25(a), we may uniquely write element in the space as

$$v = x_1 + x_2 + \ldots + x_k$$

such that $x_i \in W_i$. If $i \neq j$, we have

$$T_i T_j(v) = 0 = \delta_{ij} T_i(v)$$

by the definition of T_i's and Theorem 6.25(b). Similarly, if $i = j$, we have

$$T_i T_i(v) = x_i = \delta_{ii} T_i(v).$$

So they are the same.

6.7 The Singular Value Decomposition and the Pseudoinverse

1. (a) No. The mapping from $\mathbb{R}^2$ to $\mathbb{R}$ has no eigenvalues.

   (b) No. It's the the positive square root of the eigenvalues of $A^* A$. For example, the singular value of $2I_2$ is $2, 2$ but not eigenvalues of $(2I)^*(2I)$, which is $4, 4$.

   (c) Yes. The eigenvalue of $A^* A$ is σ^2. And the singular value of cA is the positive square root of the eigenvalue of $(cA)^*(cA) = |c|^2 A^* A$, which is $|c|^2 \sigma^2$. So the singular value of cA is $|c|\sigma$.

   (d) Yes. This is the definition.

   (e) No. For example, the singular value of $2I_2$ is $2, 2$ but not eigenvalues of $(2I)^*(2I)$, which is $4, 4$.

   (f) No. If $Ax = b$ is inconsistent, then $A^\dagger b$ could never be the solution.

   (g) Yes. The definition is well-defined.

2. For these problems, choose an orthonormal basis α, usually the standard basis, for the inner product space. Write down $A = [T]_\alpha$. Pick an orthonormal basis β such that $[T^*T]_\beta$ is diagonal. Here the order of β should follow the order of the value of its eigenvalue. The positive square roots of eigenvalues of A^*A is the singular values of A and T. Extend $T(\beta)$ to be an orthonormal basis γ for W. Then we have

$$\beta = \{v_1, v_2, \ldots, v_n\}$$

and

$$\gamma = \{u_1, u_2, \ldots, u_m\}.$$

(a) Pick α to be the standard basis. We have

$$\beta = \{(1,0), (0,1)\},$$

$$\gamma = \{\frac{1}{\sqrt{3}}(1,1,1), \frac{1}{\sqrt{2}}(0,1,-1), \frac{\sqrt{8}}{\sqrt{3}}(-\frac{1}{2}, \frac{1}{4}, \frac{1}{4})\},$$

and the singular values are $\sqrt{3}, \sqrt{2}$.

(b) Pick

$$\alpha = \{f_1 = \frac{1}{\sqrt{2}}, f_2 = \sqrt{\frac{3}{2}}x, f_3 = \sqrt{\frac{5}{8}}(3x^2 - 1)\}.$$

We have

$$\beta = \{f_3, f_1, f_2\},$$

$$\gamma = \{f_1, f_2\},$$

and the singular values are $\sqrt{45}$.

(c) Pick

$$\alpha = \{f_1 = \frac{1}{\sqrt{2\pi}}, f_2 = \frac{1}{\sqrt{\pi}}\sin x, f_3 = \frac{1}{\sqrt{\pi}}\cos x\}.$$

We have

$$\beta = \{f_2, f_3, f_1\},$$

$$\gamma = \{\frac{1}{\sqrt{5}}(2f_2 + f_3), \frac{1}{\sqrt{5}}(2f_3 - f_2), f_1\},$$

and the singular values are $\sqrt{5}, \sqrt{5}, \sqrt{4}$.

(d) Pick α to be he standard basis. We have

$$\beta = \{\frac{1}{\sqrt{3}}(1, i + 1), \sqrt{\frac{2}{3}}(1, -\frac{i+1}{2})\},$$

$$\gamma = \{\frac{1}{\sqrt{3}}(1, i + 1), \sqrt{\frac{2}{3}}(1, -\frac{i+1}{2})\},$$

and the singular values are $2, 1$.

3. Do the same to L_A as that in Exercise 6.7.2. But the α here must be the standard basis. And the matrix consisting of column vectors β and γ is V and U respectively.

(a) We have $A^*A = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}$. So its eigenvalue is $6, 0$ with eigenvectors

$$\beta = \{\frac{1}{\sqrt{2}}(1,1), \frac{1}{\sqrt{2}}(1,-1)\}.$$

Extend $L_A(\beta)$ to be an orthonormal basis

$$\gamma = \{\frac{1}{\sqrt{3}}(1,1,-1), \frac{1}{\sqrt{2}}(1,-1,0), \frac{1}{\sqrt{6}}(1,1,2)\}.$$

So we know that

$$U = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}, \Sigma = \begin{pmatrix} \sqrt{6} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, V = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

(b) We have

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \Sigma = \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix}, V = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

(c) We have

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} \frac{2}{\sqrt{10}} & 0 & 0 & -\frac{3}{\sqrt{15}} \\ \frac{1}{\sqrt{10}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{15}} \\ \frac{1}{\sqrt{10}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{15}} \\ \frac{2}{\sqrt{10}} & 0 & -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{15}} \end{pmatrix}, \Sigma = \begin{pmatrix} \sqrt{5} & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, V = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

(d) We have

$$U = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{\sqrt{2}}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \Sigma = \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & 1 \end{pmatrix}, V = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

(e) We have

$$U = \frac{1}{2}\begin{pmatrix} 1+i & -1+i \\ 1-i & 1+i \end{pmatrix}, \Sigma = \begin{pmatrix} \sqrt{6} & 0 \\ 0 & 0 \end{pmatrix}, V = \begin{pmatrix} \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{i+1}{\sqrt{6}} & \frac{-i-1}{\sqrt{3}} \end{pmatrix}.$$

(f) We have

$$U = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \end{pmatrix}, \Sigma = \begin{pmatrix} \sqrt{6} & 0 & 0 & 0 \\ 0 & \sqrt{6} & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 \end{pmatrix}, V = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

4. Find the singular value decomposition $A = U\Sigma V^*$. Then we have $W = UV^*$ and $P = V\Sigma V^*$.

   (a) We have

   $$W = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}, P = \begin{pmatrix} \sqrt{2} + \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} - \sqrt{2} \\ \frac{1}{\sqrt{2}} - \sqrt{2} & \sqrt{2} + \frac{1}{\sqrt{2}} \end{pmatrix}.$$

   (b) We have

   $$W = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, P = \begin{pmatrix} 20 & 4 & 0 \\ 4 & 20 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

5. Use the notation in Exercise 6.7.2. Then we have

$$T^\dagger(y) = \sum_{i=1}^{r} \frac{1}{\sigma_i} \langle y, u_i \rangle v_i.$$

   (a)
   $$T^\dagger(x_1, x_2, x_3) = \left(\frac{x_1 + x_2 + x_3}{3}, \frac{x_2 - x_3}{2} \right).$$

   (b)
   $$T^\dagger(a + bx + cx^2) = a\sqrt{\frac{2}{45}} f_3.$$

   (c)
   $$T^\dagger(a + b\sin x + c\cos x) = \frac{a}{2} + \frac{(2b+c)\sin x + (-b+2c)\cos x}{5}.$$

   (d)
   $$T^\dagger(z_1, z_2) = \frac{1}{2}(-z_1 + (1-i)z_2, (1+i)z_2).$$

6. Use Theorem 6.29. So we compute $A^\dagger = V\Sigma^\dagger U^*$.

   (a)
   $$A^\dagger = \frac{1}{6} \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \end{pmatrix}.$$

   (b)
   $$A^\dagger = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

(c)
$$A^\dagger = \frac{1}{5} \begin{pmatrix} 1 & -2 & 3 & 1 \\ 1 & 3 & -2 & 1 \end{pmatrix}.$$

(d)
$$A^\dagger = A^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{pmatrix}.$$

(e)
$$A^\dagger = \frac{1}{6} \begin{pmatrix} 1 - i & 1 + i \\ 1 & i \end{pmatrix}.$$

(f)
$$A^\dagger = \begin{pmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}.$$

7. Use the Lemma before Theorem 6.30. We have $Z_1 = N(T)^\perp$ and $Z_2 = R(T)$.

   (a) We have $Z_1 = \mathrm{span}\{v_1, v_2\} = V$ and $Z_2 = \mathrm{span}\{u_1, u_2\}$.
   (b) We have $Z_1 = \mathrm{span}\{v_1\}$ and $Z_2 = \mathrm{span}\{u_1\}$.
   (c) We have $Z_1 = Z_2 = V = W$.
   (d) We have $Z_1 = Z_2 = \mathbb{C}^2$.

8. If the equation is $Ax = b$, then the answer is $A^\dagger b$.

   (a) The solution is
   $$(x_1, x_2) = (\frac{1}{2}, \frac{1}{2}).$$

   (b) The solution is
   $$(x_1, x_2, x_3, x_4) = (\frac{1}{2}, 0, 1, \frac{1}{2}).$$

9. (a) Use the fact that
   $$\langle T^*(u_i), v_j \rangle = \begin{cases} \langle u_i, \sigma_j u_j \rangle & \text{if } j \le r; \\ \langle u_i, 0 \rangle & \text{if } j > r, \end{cases}$$
   $$= \delta_{ij} \sigma_j$$
   for $i \le r$. We know that $T^*(u_i) = \sigma_i v_i$. And so
   $$TT^*(u_i) = T(\sigma_i v_i) = \sigma^2 v_i$$
   for $j \le r$. Similarly, we know that, for $i > r$,
   $$\langle T^*(u_i), v_j \rangle = 0.$$
   Hence $TT^*(u_i) = 0$ when $i > r$.

(b) Let $T = L_A$ and use the previous argument.

(c) Use Theorem 6.26 and Exercise 6.7.9(a).

(d) Replace T by A.

10. Let β' and γ' be the standard bases for $\mathbb{F}^m$ and $\mathbb{F}^n$ respectly. Thus we have $[L_A]_{\beta'}^{\gamma'} = A$. Also, let

$$\beta = \{v_1, v_2, \ldots, v_m\}$$

and

$$\gamma = \{u_1, u_2, \ldots, u_n\}.$$

By Theorem 6.26 we know that $[L_A]_\beta^\gamma = \Sigma$. Apply Exercise 2.5.8 we have

$$A = [L_A]_{\beta'}^{\gamma'} = [I]_\gamma^{\gamma'}[L_A]_\beta^\gamma[I]_{\beta'}^\beta = U\Sigma V^*.$$

11. (a) Since T is normal, we have that $T^*(x) = \bar{\lambda}x$ if $T(x) = \lambda x$. Hence we know that the eigenvalues of T^*T are

$$|\lambda_1|^2, |\lambda_2|^2, \ldots, |\lambda_n|^2.$$

So we know the singular values are

$$|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|.$$

(b) Replace T by A.

12. Let $\theta_i = \frac{\lambda_i}{|\lambda_i|}$. Now we know that

$$Av_i = \lambda_i v_i = \theta_i |\lambda_i| v_i.$$

This means that $AV = U\Sigma$ and so $A = U\Sigma V^*$.

13. If A is a positive semidefinite matrix with eigenvalues λ_i's, we know that $A^* = A$ and so λ_i is real and nonnegative. Furthermoer, we have

$$A^*A(x) = A^2(x) = \lambda_i^2 x.$$

Since $\lambda_i \geq 0$, we have $\sqrt{\lambda_i^2} = \lambda_i$. Hence the eigenvalues of A are the singular values of A.

14. Consider

$$A^2 = A^*A = V\Sigma U^*U\Sigma V^* = V\Sigma^2 V^* = (V\Sigma V^*)^2.$$

Both of A and $V\Sigma V^*$ are positive definite, we know that $A = V\Sigma V^*$ by Exercise 6.4.17(d). So we know

$$V\Sigma V^* = U\Sigma V^*.$$

Since A is positive definite, we know Σ is invertible. Also, V is invertible. Hence we get $U = V$ finally.

15. (a) Use the fact that
$$A^*A = P^*W^*WP = P^2$$

and
$$AA^* = WPP^*W^* = WP^2W^*.$$

So $A^*A = AA^*$ if and only if $P^2 = WP^2W^*$, which is equivalent to $WP^2 = P^2W$.

(b) By the previous argument we have A is normal if and only if
$$P^2 = WP^2W^* = (WPW^*)^2.$$

Since P and WPW^* are both positive semidifinite, again by Exercise 6.4.17(d), we have the condition is equivalent to $P = WPW^*$ or $PW = WP$.

16. Use the singular value decomposition $A = U\Sigma V^*$. Let $P = U\Sigma U^*$ and $W = UV^*$. Then we have $A = PW$ and P is positive semidefinite and W is unitary.

17. (a) We calculate $(UT)^\dagger$ and $U^\dagger$, $T^\dagger$ separately. First we have
$$UT(x_1, x_2) = U(x_1, 0) = (x_1, 0) = T^\dagger.$$

So we compute $T^\dagger$ directly. We have $N(T)$ is the y-axis. So $N(T)^\perp$ is the x-axis. Since we have
$$T(1, 0) = (1, 0),$$

we know that
$$T^\dagger(1, 0) = (1, 0)$$

and
$$T^\dagger(0, 1) = (0, 0).$$

Hence we have
$$(UT)^\dagger(x_1, x_2) = T^\dagger(x_1, x_2) = x_1 T^\dagger(1, 0) + x_2 T^\dagger(0, 1) = x_1 T^\dagger(1, 0) = (x_1, 0).$$

On the other hand, we also have $N(U)$ is the line span$\{(1, -1)\}$. So $N(U)^\perp$ is the line span$\{(1, 1)\}$. Since we have
$$U(1, 1) = (2, 0),$$

we know that
$$U^\dagger(1, 0) = \frac{1}{2}(1, 1)$$

and
$$U^\dagger(0, 1) = (0, 0).$$

Hence we have

$$U^\dagger(x_1, x_2) = x_1 U^\dagger(1,0) + x_2 U^\dagger(0,1) = (\frac{x_1}{2}, \frac{x_1}{2}).$$

Finally we have

$$T^\dagger U^\dagger(x_1, x_2) = T^\dagger(\frac{x_1}{2}, \frac{x_1}{2}) = (\frac{x_1}{2}, 0) \neq (UT)^\dagger(x_1, x_2).$$

(b) Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. By the previous argument, we have $A^\dagger = \begin{pmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \end{pmatrix}$ and $B^\dagger = B$. Also, we have $AB = B$ and so $(AB)^\dagger = B^\dagger = B$.

18. (a) Observe that if $A = U\Sigma V^*$ is a singular value decomposition of A, then $GA = (GU)\Sigma V^*$ is a single value decomposition of GA. So we have
$$(GA)^\dagger = V\Sigma^\dagger(GU)^* = A^\dagger G^*.$$

(b) Observe that if $A = U\Sigma V^*$ is a singular value decomposition of A, then $AH = U\Sigma V^* H^{**} = U\Sigma(H^*V)^*$ is a single value decomposition of AH. So we have
$$(AH)^\dagger = H^*V\Sigma^\dagger U^* = H^* A^\dagger.$$

19. (a) The nonzero singular values of A are the positive square roots of the nonzero eigenvalues of A^*A. But the eigenvalues of A^*A and that of AA^* are the same by Exercise 6.7.9(c). Hence we know that the singular value decomposition of A and that of A^* are the same.
Also, we have
$$(A^t)^* A^t = \overline{AA}^* = \overline{AA^*}.$$

Since AA^* is self-adjoint, its eigenvalues are always real. We get that if $AA^*(x) = \lambda x$, then we have
$$(A^t)^* A^t(\overline{x}) = \overline{AA^*(x)} = \overline{\lambda x} = \lambda \overline{x},$$

here $\overline{A}$ means the matrix consisting of the conjugate of the entries of A. Hence the singular value of A^t and that of A^* are all the same.

(b) Let $A = U\Sigma V^*$ be a singular value decomposition of A. Then we have $A^* = V\Sigma^* U^*$. So
$$(A^*)^\dagger = U(\Sigma^*)^\dagger V^* = U(\Sigma^\dagger)^* V^* = (A^\dagger)^*.$$

(c) Let $A = U\Sigma V^*$ be a singular value decomposition of A. Then we have $A^t = (V^*)^t \Sigma^t U^t$. So
$$(A^t)^\dagger = (U^t)^* (\Sigma^t)^\dagger V^t = (U^*)^t (\Sigma^\dagger)^t V^t = (A^\dagger)^t.$$

20. Let $A = U\Sigma V^*$ be a singular value decomposition of A. Then we have

$$O = A^2 = U\Sigma V^* U\Sigma V,$$

which means

$$\Sigma V^* U\Sigma = O$$

since U and V are invertible. Now let $\{\sigma_i\}_{i=1}^r$ is the set of those singular values of A. Denote D to be the diagonal matrix with $D_{ii} = \frac{1}{\sigma^2}$ if $i \leq r$ while $D_{ii} = 1$ if $i > r$. Then we have $\Sigma D = D\Sigma = \Sigma^\dagger$. This means that

$$\Sigma^\dagger V^* U\Sigma^\dagger = D\Sigma V^* U\Sigma D = DOD = O.$$

Now we have

$$(A^\dagger)^2 = (V\Sigma^\dagger U^*)^2 = V\Sigma^\dagger U^* V\Sigma^\dagger U^*$$

$$= V(\Sigma^\dagger V^* U\Sigma^\dagger)^* U^* = VOU^* = O.$$

21. Here we use the notation in 6.26.

  (a) Compute

  $$TT^\dagger T(v_i) = TT^\dagger(\sigma_i u_i) = T(v_i)$$

  if $i \leq r$, while

  $$TT^\dagger T(v_i) = TT^\dagger(0) = 0 = T(v_i)$$

  if $i > r$.

  (b) Compute

  $$T^\dagger TT^\dagger(u_i) = T^\dagger T(\frac{1}{\sigma_i} v_i) = T^\dagger(u_i)$$

  if $i \leq r$, while

  $$T^\dagger TT^\dagger(u_i) = T^\dagger T(0) = 0 = T^\dagger(u_i)$$

  if $i > r$.

  (c) Pick an orthonormal basis α. Let A be the matrix $[T]_\alpha$ and $A = U\Sigma V^*$ be a singular value decomposition of A. Then we have

  $$(A^\dagger A)^* = (V\Sigma^\dagger U^* U\Sigma V^*)^* = (V\Sigma^\dagger \Sigma V^*)^* = A^\dagger A$$

  and

  $$(AA^\dagger)^* = (U\Sigma V^* V\Sigma^\dagger U^*)^* = (U\Sigma\Sigma^\dagger U^*)^* = AA^\dagger.$$

22. Observe that UT is the orthogonal projection on $R(UT)$ by Theorem 6.24 since it's self-adjoint and $UTUT = UT$. We have that $R(UT) = R(T^*U^*) \subset R(T^*)$. Also, since

$$UTT^*(x) = T^*U^*T^*(x) = T^*(x),$$

216

we have $R(T^*) \subset R(UT)$ and hence $R(T^*) = R(UT)$. This means UT and $T^\dagger T$ are both orthogonal projections on $R(T^*) = R(UT)$. By the uniqueness of orthogonal projections, we have $UT = T^\dagger T$.

Next, observe that TU is the orthogonal projection on $R(TU)$ by Theorem 6.24 since it's self-adjoint and $TUTU = TU$. We have that $R(TU) \subset R(T)$. Also, since

$$TUT(x) = T(x),$$

we have $R(T) \subset R(TU)$. By the same reason, we have $TU = TT^\dagger$ and they are the orthogonal projection on $R(T) = R(TU)$.

Finally, since we have $TU - TT^\dagger = T_0$, we may write is as

$$T(U - T^\dagger) = T_0.$$

We want to claim that $R(U - T^\dagger) \cap N(T) = \{0\}$ to deduce that $U - T^\dagger = T_0$. Observe that $R(T^\dagger) = N(T)^\perp = R(T^*)$. Also, we have $R(U) \subset R(T^*)$ otherwise we may pick $x \in W$ such that $U(x) \in R(U) \backslash R(T^*)$ and get the contradiction that

$$0 \neq U(x) = UTU(x) = 0$$

since UT is the orthogonal projection on $R(T^*)$. Now we already have $R(U - T^\dagger) \subset R(T^*) = N(T)$. Hence the claim

$$R(U - T^\dagger) \cap N(T) = \{0\}$$

holds. This means that $T(U - T^\dagger)(x) = 0$ only if $(U - T^\dagger)(x) = 0$. Since we have $T(U - T^\dagger) = T_0$, now we know that actually $U - T^\dagger = T_0$ and hence $U = T^\dagger$.

23. Replace T by A.

24. Replace T by A.

25. (a) By Exercise 6.3.13 T^*T is invertible. Let $U = (T^*T)^{-1}T^*$. Check that
$$TUT = T(T^*T)^{-1}T^*T = T$$
and
$$UTU = (T^*T)^{-1}T^*T(T^*T)^{-1}T^* = U.$$
Also, both $TU = T(T^*T)^{-1}T^*$ and $UT = (T^*T)^{-1}T^*T = I$ are self-adjoint. Apply Exercise 6.7.21 and get the result.

(b) By Exercise 6.3.13 TT^* is invertible. Let $U = T^*(TT^*)^{-1}$. Check that
$$TUT = TT^*(TT^*)^{-1}T = T$$
and
$$UTU = T^*(TT^*)^{-1}TT^*(TT^*)^{-1} = U.$$
Also, both $TU = TT^*(TT^*)^{-1} = I$ and $UT = T^*(TT^*)^{-1}T$ are self-adjoint. Apply Exercise 6.7.21 and get the result.

26. By Theorem 6.26, we know $[T]_{\beta'}^{\gamma'} = \Sigma$ for some orthonormal bases β' and γ', where Σ is the matrix in Theorem 6.27. In this case we know that

$$([T]_{\beta'}^{\gamma'})^\dagger = \Sigma^\dagger = [T^\dagger]_{\gamma'}^{\beta'}.$$

Now for other orthonormal bases β and γ. We know that

$$[T]_\beta^\gamma = [I]_{\gamma'}^\gamma \Sigma [I]_\beta^{\beta'}$$

is a singular value decomposition of $[T]_\beta^\gamma$ since both $[I]_{\gamma'}^\gamma$ and $[I]_\beta^{\beta'}$ are unitary by the fact that all of them are orthonormal. Hence we have

$$([T]_\beta^\gamma)^\dagger = [I]_{\beta'}^\beta \Sigma^\dagger [I]_\gamma^{\gamma'} = [I]_{\beta'}^\beta [T^\dagger]_{\gamma'}^{\beta'} [I]_\gamma^{\gamma'} = [T^\dagger]_\gamma^\beta.$$

27. Use the notation in Theorem 6.26. By the definition of $T^\dagger$, we have

$$TT^\dagger(x) = LL^{-1}(x) = x.$$

Also, if $x \in R(T)^\dagger$, again by the definition of $T^\dagger$, we have $TT^\dagger(x) = 0$. Hence it's the orthogonal projection of V on $R(T)$.

6.8 Bilinear and Quadratic Forms

1. (a) No. A quadratic form is a function of one variable. But a bilinear form is a function of two variables.

   (b) No. We have $4I = (2I)^t I (2I)$. But $4I$ and $2I$ has different eigenvalues.

   (c) Yes. This is Theorem 6.34.

   (d) No. See Example 5 of this section. The matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is a counterexample when $\mathbb{F} = \mathbb{Z}_2$.

   (e) Yes. Let H_1 and H_2 be two symmetric bilinear forms. We have

   $$(H_1 + H_2)(x, y) = H_1(x, y) + H_2(x, y)$$

   $$= H_1(y, x) + H_2(y, x) = (H_1 + H_2)(y, x).$$

   (f) No. The bilinear forms

   $$H_1(x, y) = x^t \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} y, H_2(x, y) = x^t \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} y$$

   have matrix representations $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ respectively with the standard basis. But both of their characteristic polynomials are $(1 - t)^2$.

(g) No. We must have $H(0,0) = 0$ since $H(x,0)$ is a linear function of x.

(h) No. It's $n^2 \neq 2n$ for $n \neq 2$ by the Corollary 1 after Theorem 6.32.

(i) Yes. Pick a nonzero element $u \in V$ arbitrarily. If $H(x,u) = 0$, then we have $y = u$. Otherwise pick another nonzero element $v \in V$ such that $\{u,v\}$ is independent. Thus we have

$$y = H(x,v)u - H(x,u)v \neq 0.$$

But we have

$$H(x,y) = H(x,v)H(x,u) - H(x,u)H(x,v) = 0.$$

(j) No. It needs one more condition that H is symmetric. For example, the matrix $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has its congruent matrix

$$Q^t \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} Q = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ac & bc \\ ad & bd \end{pmatrix}.$$

If that congruent matrix is diagonal, we should have $bc = ad = 0$. If $a = b = 0$ or $a = c = 0$, then Q is not invertible. Similarly, it cannot be $d = c = 0$ or $d = b = 0$. So this bilinaer form is not even diagonalizable.

2. The property 1 comes from the definition of a bilinear form. The property 2 comes from that $L_x(0) = R_x(0) = 0$. The property 3 comes from the computation

$$H(x + y, z + w) = H(x, z + w) + H(y, z + w)$$

$$= H(x, z) + H(x, w) + H(y, z) + H(y, w).$$

Finally, since the conditions in the definition of a bilinear form is symmetric for the first and the second component. So we get the property 4.

3. (a) Check that

$$(H_1 + H_2)(ax_1 + x_2, y) = H_1(ax_1 + x_2, y) + H_2(ax_1 + x_2, y)$$

$$= aH_1(x_1, y) + H_1(x_2, y) + aH_2(x_1, y) + H_2(x_2, y)$$

$$= a(H_1 + H_2)(x_1, y) + (H_1 + H_2)(x_2, y)$$

and

$$(H_1 + H_2)(x, ay_1 + y_2) = H_1(x, ay_1 + y_2) + H_2(x, ay_1 + y_2)$$

$$= aH_1(x, y_1) + H_1(x, y_2) + aH_2(x, y_1) + H_2(x, y_2)$$

$$= a(H_1 + H_2)(x, y_1) + (H_1 + H_2)(x, y_2).$$

(b) Check that
$$cH(ax_1 + x_2, y) = caH(x_1, y) + cH(x_2, y)$$
$$= acH(x_1, y) + cH(x_2, y)$$
and
$$cH(x, ay_1 + y_2) = caH(x, y_1) + cH(x, y_2)$$
$$= acH(x, y_1) + cH(x, y_2).$$

(c) Pick $H_0(x, y) = 0$ as the zero element and check the condition for a vector space.

4. (a) Yes. The form $f \times g$ is bilinear and the integral operator is linear.
   (b) No. If $J(x, y) \neq 0$ for some x and y, then we have
   $$H(cx, y) = [J(cx, y)]^2 = c^2[J(x, y)]^2 \neq cH(x, y).$$

   (c) No. We have
   $$H(2, 1) = 4 \neq 2H(1, 1) = 6.$$

   (d) Yes. The determinant function is an n-linear function and now n is 2.

   (e) Yes. When the field if $\mathbb{R}$, the inner product function is a bilinear form.

   (f) No. It fails when $\mathbb{F} = \mathbb{C}$. If we pick $V = \mathbb{C}$ and choose the standard inner product. Thus we have
   $$H(1, i) = \langle 1, i \rangle = -1 \neq iH(1, 1) = i.$$

5. See the definition of the matrix representation.

   (a) It's a bilinear form since
   $$H\left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\right) = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}^t \begin{pmatrix} 1 & -2 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$
   The matrix representation is
   $$\begin{pmatrix} 0 & 2 & -2 \\ 2 & 0 & -2 \\ 1 & 1 & 0 \end{pmatrix}.$$

   (b) It's a bilinear form since
   $$H\left(\begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}, \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix}\right) = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}^t \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}.$$
   The matrix above is the matrix representation with respect to the standard basis.

(c) Let
$$f = a_1 \cos t + a_2 \sin t + a_3 \cos 2t + a_4 \sin 2t$$
and
$$g = b_1 \cos t + b_2 \sin t + b_3 \cos 2t + b_4 \sin 2t.$$

We compute that
$$H(f,g) = (a_2 + 2a_4)(-b_1 - 4b_3)$$

$$= \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}^t \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 0 & -4 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & -4 & 0 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}.$$

Hence it's a bilinear form. And the matrix is the matrix representation.

6. We have
$$H\left(\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\right) = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}^t \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Hence we find the matrix A. And the form $x^t A y$ is a bilinear form.

7. (a) Check that
$$\widehat{T}(H)(ax_1, x_2, y) = H(T(ax_1 + x_2), T(y))$$

$$= H(aT(x_1) + T(x_2), T(y)) = aH(T(x_1), T(y)) + H(T(x_2), T(y))$$
$$= a\widehat{T}(H)(x_1, y) + \widehat{T}(H)(x_2, y)$$

and
$$\widehat{T}(H)(x, ay_1 + y_2) = H(T(x, T(ay_1 + y_2)))$$
$$= H(T(x), aT(y_1) + T(y_2)) = aH(T(x), T(y_1)) + H(T(x), T(y_2))$$
$$= a\widehat{T}(H)(x, y_1) + \widehat{T}(H)(x, y_2).$$

(b) Check that
$$\widehat{T}(cH_1 + H_2)(x, y) = (cH_1 + H_2)(T(x), T(y))$$

$$= cH_1(T(x), T(y)) + H_2(T(x), T(y))$$
$$= [c\widehat{T}(H_1) + \widehat{T}(H)](x, y).$$

(c) Suppose T is injective and surjective. If H is an nonzero bilinear form with $H(x_1, y_1) \neq 0$ for some $x_1, y_1 \in W$ and $\widehat{T}(H)$ is the zero bilinear form, we may find $x_0, y_0 \in V$ such that $T(x_0) = x_1$ and $T(y_0) = y_1$ since T is surjective. Thus we'll have
$$0 = \widehat{T}(H)(x_0, x_1) = H(x, y) \neq 0,$$

a contradiction. This means that $\widehat{T}$ is injective. On the other hand, since T is an isomorphism, the inverse of T exists. Then for each $H \in \mathcal{B}(V)$, we can define

$$H_0(x, y) := H_0(T^{-1}(x), T^{-1}(y))$$

such that

$$\widehat{T}(H_0) = H.$$

8. (a) Let $\beta = \{v_i\}$. We know that $(\psi_\beta(H))_{ij} = H(v_i, v_j)$. So we have

$$(\psi_\beta(cH_1 + H_2))_{ij} = (cH_1 + H_2)(v_i, v_j)$$

$$= cH_1(v_i, v_j) + H_2(v_i, v_j) = c(\psi_\beta(H_1))_{ij} + (\psi_\beta(H_2))_{ij}.$$

(b) The form $H'(u, v) := u^t A v$ is a bilinear form when $u, v \in \mathbb{F}^n$. We know that ϕ_β is an isomorphism from V to $\mathbb{F}^n$. This means that

$$H = \widehat{\phi_\beta^{-1}}(H')$$

is a bilinear form by Exercise 6.8.7.

(c) Let $\beta = \{v_i\}$. And let

$$x = \sum_{i=1}^n a_i v_i, y = \sum_{i=1}^n b_i v_i.$$

Thus we have

$$H(x, y) = H\left(\sum_{i=1}^n a_i v_i, \sum_{i=1}^n b_i v_i\right)$$

$$\sum_{i,j} a_i b_i H(v_i, v_j) = [\phi_\beta(x)]^t A[\phi_\beta(y)].$$

9. (a) It comes from the fact that $\dim(M_{n \times n}(\mathbb{F})) = n^2$.

(b) Let E_{ij} be the matrix whose ij-entry is 1 and other entries are zero. Then we know that $\{E_{ij}\}_{ij}$ is a basis in $M_{n \times n}(\mathbb{F})$. Since ψ_β is an isomorphism, the set $\psi_\beta^{-1}(\{E_{ij}\}_{ij})$ is a basis for $\mathcal{B}(V)$.

10. The necessity comes from Exercise 6.8.8(c). For the sufficiency, we know that

$$(\psi_\beta(H))_{ij} = H(v_i, v_j) = e_i^t A e_j = A_{ij},$$

where v_i, v_j are elements in β and e_i, e_j are the elements in the standard basis in $\mathbb{F}^n$.

11. Pick β to be the standard basis and apply the Corollary 3 after Theorem 6.32. Thus we have $[\phi_\beta(x)] = x$.

12. Prove the three conditions.

reflexivity We have A is congruent to A since $A = I^t A I$.

symmetry If A is congruent to B, we have $B = Q^t A Q$ for some invertible matrix Q. Hence we know that B is congruent to A since $A = (Q^{-1})^t A Q^{-1}$.

transitivity If A is congruent to B and B is congruent to C, we have $B = Q^t A Q$ and $C = P^t B P$. Thus we know that A is congruent to C since $C = (QP)^t A (QP)$.

13. (a) If x is an element in V, then $\phi_\gamma(x)$ and $\phi_\beta(x)$ are the γ-coordinates and the β-coordinates respectly. By the definition of Q, we have

$$\phi_\beta(x) = L_Q \phi_\gamma(x)$$

for all $x \in V$.

(b) By the Corollary 2 after Theorem 6.32, we know that

$$H(x,y) = [\phi_\gamma(x)]^t \psi_\gamma(H)[\phi_\gamma(y)] = [\phi_\beta(x)]^t \psi_\beta(H)[\phi_\beta(y)].$$

By the previous argument we know that

$$[\phi_\gamma(x)]^t Q^t \psi_\beta(H) Q [\phi_\gamma(y)] = [\phi_\gamma(x)]^t \psi_\gamma(H)[\phi_\gamma(y)],$$

where Q is the change of coordinate matrix changing γ-coordinates to β-coordinates. Again, by the Corrolary 2 after Theorem 6.32 we know the matrix $Q^t \psi_\beta(H) Q$ must be the matrix $\psi_\gamma(H)$. Hence they are congruent.

14. Since they are congruent, we have

$$Q^t \psi_\beta(H) Q = \psi_\gamma(H)$$

for some invertible matrix Q. But invertible matrix will preserve the rank, so we know their rank are the same.

15. (a) If A is a square diagonal matrix, then we have $A_{ij} = A_{ji} = 0$.

(b) If A is a matrix congruent to a diagonal matrix B, then we have $B = Q^t A Q$ and $A = (Q^{-1})^t B Q^{-1}$. This means A is symmetric since

$$A^t = (Q^{-1})^t B^t Q^{-1} = (Q^{-1})^t B Q^{-1} = A.$$

(c) Say α to be the standard basis and β to be the basis in Theorem 6.35. Let $H = \psi_\alpha^{-1}(A)$ be the bilinear form whose matrix representation is A. Thus we know that $\psi_\alpha(H) = A$ and $\psi_\beta(H)$ are congruent. Also, by Theorem 6.35 we know that $\psi_\beta(H)$ is diagonal.

16. If $K(x) = H(x,x)$, then we have

$$K(x+y) = H(x+y, x+y) = H(x,x) + 2H(x,y) + H(y,y)$$

$$= K(x) + 2H(x, y) + K(y).$$

If $\mathbb{F}$ is not of characteristic two, we get the formula

$$H(x, y) = \frac{1}{2}[K(x + y) - K(x) - K(y)].$$

17. Use the formula given in the previous exercise to find H. To diagonalize it, we may use the method in the paragraph after Theorem 6.35. Here the notation α is the standard basis in the corresponding vector spaces.

   (a) We have

   $$H\left(\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\right) = \frac{1}{2}[K\begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \end{pmatrix} - K\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} - K\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}]$$

   $$= -2a_1b_1 + 2a_1b_2 + 2a_2b_1 + a_2b_2.$$

   Also, we have $\phi_\alpha = \begin{pmatrix} -2 & 2 \\ 2 & 1 \end{pmatrix}$. Hence we know that

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

   So the basis β could be

   $$\{(1, 0), (1, 1)\}.$$

   (b) We have

   $$H\left(\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\right) = 7a_1b_1 - 4a_1b_2 - 4a_2b_1 + a_2b_2$$

   and

   $$\beta = \{(1, 0), (\frac{4}{7}, 1)\}.$$

   (c) We have

   $$H\left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\right) = 3a_1b_1 + 3a_2b_2 + 3a_3b_3 - a_1b_3 - a_3b_1.$$

   and

   $$\beta = \{(1, 0, 0), (0, 1, 0), (\frac{1}{3}, 0, 1)\}.$$

18. As what we did in the previous exercise, we set

$$K\begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix} = 3t_1^2 + 3t_2^2 + 3t_3^2 - 2t_1t_3$$

and find

$$H\left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\right) = 3a_1b_1 + 3a_2b_2 + 3a_3b_3 - a_1b_3 - a_3b_1$$

such that $H(x,x) = K(x)$ and H is a bilinear form. This means that

$$K\begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix} = \begin{pmatrix} t_1 & t_2 & t_3 \end{pmatrix} \begin{pmatrix} 3 & 0 & -1 \\ 0 & 3 & 0 \\ -1 & 0 & 3 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}$$

$$= \begin{pmatrix} t_1 & t_2 & t_3 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}.$$

Note that here we may diagonalize it in sence of eigenvectors. Thus we pick

$$\beta = \{(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}), (0, 1, 0).\}$$

And take

$$\begin{pmatrix} t_1' \\ t_2' \\ t_3' \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}.$$

Thus we have

$$3t_1^2 + 3t_2^2 + 3t_3^2 - 2t_1t_= \begin{pmatrix} t_1' & t_2' & t_3' \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} t_1' \\ t_2' \\ t_3' \end{pmatrix}$$

$$= 4(t_1')^2 + 2(t_2')^2 + 3(t_3')^2.$$

Hence the original equality is that

$$4(t_1')^2 + 2(t_2')^2 + 3(t_3')^2 + l.o.t = 0,$$

where $l.o.t$ means some lower order terms. Hence $\mathcal{S}$ is a ellipsoid.

19. Here we use the noation in the proof of Theorem 6.37. Also, the equation

$$\sum_{i=1}^{n} (\frac{1}{2}\lambda_i - \epsilon)s_i^2 < f(x) < \sum_{i=1}^{n} (\frac{1}{2}\lambda_i + \epsilon)s_i^2$$

is helpful.

(a) Since $0 < \text{rank}(A) < n$ and A has no negative eigenvalues, we could find a positive eigenvalue λ_i of A. Then take $x = s_i v_i$. Then we'll have that

$$f(x) > (\frac{1}{2}\lambda_i - \epsilon)s_i^2 > 0 = f(0).$$

We may pick s_i arbitrarily small such that $\|x - p\|$ could be arbitrarily small. Hence f has no local maximum at p.

(b) Since $0 < \text{rank}(A) < n$ and A has no positive eigenvalues, we could find a negative eigenvalue λ_i of A. Then take $x = s_i v_i$. Then we'll have that

$$f(x) < (\frac{1}{2}\lambda_i + \epsilon)s_i^2 < 0 = f(0).$$

We may pick s_i arbitrarily small such that $\|x-p\|$ could be arbitrarily small. Hence f has no local minimum at p.

20. Observe that D is the determinant of the Hessian matrix A. Here we denote λ_1, λ_2 to be the two eigenvalues of A, which exist since A is real symmetric.

(a) If $D > 0$, we know that λ_1 and λ_2 could not be zero. Since $\frac{\partial^2 f(p)}{\partial t_1^2} > 0$, we have $\frac{\partial^2 f(p)}{\partial t_2^2} > 0$ otherwise we'll have $D \leq 0$. Hence the trace of A is positive. Thus we have

$$\lambda_1 + \lambda_2 = -\text{tr}(A) < 0$$

and

$$\lambda_1 \lambda_2 = D > 0.$$

This means that both of them are negative. Hence p is a local minimum by the Second Derivative Test.

(b) If $D < 0$, we know that λ_1 and λ_2 could not be zero. Since $\frac{\partial^2 f(p)}{\partial t_1^2} < 0$, we have $\frac{\partial^2 f(p)}{\partial t_2^2} < 0$ otherwise we'll have $D \geq 0$. Hence the trace of A is negative. Thus we have

$$\lambda_1 + \lambda_2 = -\text{tr}(A) > 0$$

and

$$\lambda_1 \lambda_2 = D < 0.$$

This means that both of them are positive. Hence p is a local maximum by the Second Derivative Test.

(c) If $D < 0$, we know that λ_1 and λ_2 could not be zero. Also, we have

$$\lambda_1 \lambda_2 = D < 0.$$

This means that they cannot be both positive or both negative. Again, by the Second Derivative Test, it's a saddle point.

(d) If $D = 0$, then one of λ_1 and λ_2 should be zero. Apply the Second Derivative Test.

21. As Hint, we know that

$$E^t A = (A^t E)^t.$$

That is, do the same column operation on A^t. This means that do the same row operation on A.

22. See the paragraph after Theorem 6.35.

(a) Take

$$\begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix} A \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -7 \end{pmatrix}.$$

(b) Take

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

(c) Take

$$\begin{pmatrix} 1 & -\frac{1}{4} & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} A \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{4} & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{19}{4} & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

23. Since each permutation could be decomposed into several 2-cycle, interchanging two elements, we may just prove the statement when the permutation is 2-cycle. Let A be a diagonal matrix and B be the diagonal matrix obtained from A by interchanging the ii-entry and the jj-entry. Take E be the elementary matrix interchaning the i-th and the j-th row. Then we have E is symmetric and $EAE = B$.

24. (a) Compute that

$$H(ax_1 + x_2, y) = \langle ax_1 + x_2, T(y) \rangle$$

$$= a\langle x_1, T(y) \rangle + \langle x_2, T(y) \rangle$$

$$= aH(x_1, y) + H(x_2, y)$$

and

$$H(x, ay_1 + y_2) = \langle x, T(ay_1 + y_2) \rangle = \langle x, aT(x_1) + T(x_2) \rangle$$

$$= a\langle x, T(y_1) \rangle + \langle x, T(y_2) \rangle$$

$$= aH(x, y_1) + H(x, y_2).$$

(b) Compute that

$$H(y, x) = \langle y, T(x) \rangle$$

$$= \langle T(x), y \rangle = \langle x, T^*(y) \rangle.$$

The value equal to $H(x, y) = \langle x, T(y) \rangle$ for all x and y if and only if $T = T^*$.

(c) By Exercise 6.4.22 the operator T must be a positive semidifinite operator.

(d) It fail since

$$H(x, iy) = \langle x, T(iy) \rangle = \langle x, iT(y) \rangle$$

$$= -i\langle x, T(y) \rangle \neq iH(x, y)$$

in genereal.

25. Let $A = \psi_\beta(H)$ for some orthonormal basis β. And let T be the operator such that $[T]_\beta = A$. By Exercise 6.8.5 we have

$$H(x,y) = [\phi_\beta(x)]^t A[\phi_\beta(y)] = [\phi_\beta(x)]^t [T]_\beta[\phi_\beta(y)] = [\phi_\beta(x)]^t [\phi_\beta(T(y))].$$

Also, by Parseval's Identity in Exercise 6.2.15 we know that

$$\langle x, T(y)\rangle = \sum_{i=1}^{n} \langle x, v_i\rangle\langle T(y), v_i\rangle = [\phi_\beta(x)]^t [\phi_\beta(T(y))]$$

since β is orthonormal.

26. Use the Corollary 2 after Theorem 6.38. Let p, q be the number of positive and negative eigenvalues respectly. Then we have

$$p + q \leq n.$$

Hence we have

$$\binom{3+n-1}{n} = \binom{n+2}{2} = \frac{(n+2)(n+1)}{2}$$

possibilities.

6.9 Einstein's Special Theory of Relativity

1. **(b)** It comes from that $T_v(e_i) = e_i$ for $i = 2, 3$.

   (c) By the axiom (R4), we know that

   $$T_v\begin{pmatrix} a \\ 0 \\ 0 \\ d \end{pmatrix} = \begin{pmatrix} a' \\ 0 \\ 0 \\ d' \end{pmatrix}.$$

   (d) We compute that, for $i = 2, 3$ and $j = 1, 4$,

   $$\langle T_v^*(e_i), e_j\rangle = \langle e_i, T_v(e_j)\rangle = 0$$

   by the fact that $T_v(e_j) \in \text{span}(\{e_1, e_4\})$. Hence we know that $\text{span}(\{e_2, e_3\})$ is T_v^*-invariant.

   On the other hand, we compute that, for $i = 2, 3$ and $j = 1, 4$,

   $$\langle T_v^*(e_j), e_i\rangle = \langle e_j, T_v(e_i)\rangle = \langle e_j, e_i\rangle = 0.$$

   So $\text{span}(\{e_1, e_4\})$ is T_v^*-invariant.

2. We already have that
   $$\langle T_v^* L_A T_v(w), w\rangle = 0$$
   if $\langle L_A(w), w\rangle = 0$ for $w \in \mathbb{R}^4$ whose fourth entry is nonnegative. Now if we have $\langle L_A(w), w\rangle = 0$ for some w is a vector in $\mathbb{R}^4$ whose fourth entry is negative, then we have

   $$0 = \langle T_v^* L_A T_v(-w), -w\rangle = (-1)^2 \langle T_v^* L_A T_v(w), w\rangle = \langle T_v^* L_A T_v(w), w\rangle.$$

3. (a) The set $\{w_1, w_2\}$ is linearly independent by definition. Also, we have $w_1 = e_1 + e_4$ and $w_2 = e_1 - e_4$ are both elements in $\text{span}(\{e_1, e_4\})$. Hence it's a basis for $\text{span}(\{e_1, e_4\})$. Naturally, it's orthogonal since $\langle w_1, w_2 \rangle = 0$.

   (b) For brevity, we write $W = \text{span}(\{e_1, e_4\})$. We have $T_v(W) \subset W$ and $T_v^*(W) \subset W$ by Theorem 6.39. Also, W is L_A-invariant if we directly check it. Hence W is $T_v^* L_A T_v$-invariant.

4. We know that $B_v^* A B_v = [T_v^* L_A T_v]_\beta$ and $[L_A]_\beta$. So (a) and (b) in the Corollary is equivalent. We only prove (a) by the steps given by Hints. For brevity, we write $U = T_v^* L_A T_v$ and $C = B_v^* A B_v$.

   (a) We have $U(e_i) = e_i$ for $i = 2, 3$ by Theorem 6.39. By Theorem 6.41 we have

   $$\begin{cases} U(e_1) + U(e_4) = U(e_1 + e_4) = U(w_1) = aw_2 = ae_1 - ae_4, \\ U(e_1) - U(e_4) = U(e_1 - e_4) = U(w_2) = bw_1 = be_1 + be_2. \end{cases}$$

   Solving this system of equations we get that

   $$\begin{cases} U(e_1) = pe_1 - qe_4, \\ U(e_4) = qe_1 - pe_2, \end{cases}$$

   where $p = \frac{a+b}{2}$ and $q = \frac{a-b}{2}$. Write down the matrix representation of U and get the result.

   (b) Since C is self-adjoint, we know that $q = -q$ and so $q = 0$.

   (c) Let $w = e_2 + e_4$. Then we know that $\langle L_A(w), w \rangle = 0$. Now we calculate that
   $$U(w) = U(e_2 + e_4) = e_2 - pe_4.$$

   By Theorem 6.40 we know that

   $$\langle U(w), w \rangle = \langle e_2 - pe_4, e_2 + e_4 \rangle = 1 - p = 0.$$

   Hence we must have $p = 0$.

5. We only know that
   $$T_v \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -vt'' \\ 0 \\ 0 \\ t'' \end{pmatrix}$$

   for some $t'' > 0$. Compute that

   $$\left\langle T_v^* L_A T_v \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle = \left\langle L_A \begin{pmatrix} -vt'' \\ 0 \\ 0 \\ t'' \end{pmatrix}, \begin{pmatrix} -vt'' \\ 0 \\ 0 \\ t'' \end{pmatrix} \right\rangle = (t'')^2(v^2 - 1)$$

and

$$\langle T_v^* L_A T_v \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \rangle = \langle L_A \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \rangle = -1$$

by Theorem 6.41. Hence we have $(t'')^2(v^2 - 1) = -1$ and so $t'' = \frac{1}{\sqrt{1-v^2}}$.

6. Note that if S_2 is moving past S_1 at a velocity $v > 0$ as measured on S, the T_v is the transformation from space-time coordinates of S_1 to that of S_2. So now we have

$$\begin{cases} T_{v_1} : S \to S', \\ T_{v_2} : S' \to S'', \\ T_{v_3} : S \to S''. \end{cases}$$

The given condition says that $T_{v_3} = T_{v_2} T_{v_3}$. This means that

$$B_{v_2} B_{v_1} = \begin{pmatrix} \frac{1}{\sqrt{1-v_2^2}} & 0 & 0 & \frac{-v_2}{\sqrt{1-v_2^2}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{-v_2}{\sqrt{1-v_2^2}} & 0 & 0 & \frac{1}{\sqrt{1-v_2^2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{1-v_1^2}} & 0 & 0 & \frac{-v_1}{\sqrt{1-v_1^2}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{-v_1}{\sqrt{1-v_1^2}} & 0 & 0 & \frac{1}{\sqrt{1-v_1^2}} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1+v_2 v_1}{\sqrt{(1-v_2^2)(1-v_1^2)}} & 0 & 0 & \frac{-v_2-v_1}{\sqrt{(1-v_2^2)(1-v_1^2)}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{v_2-v_1}{\sqrt{(1-v_2^2)(1-v_1^2)}} & 0 & 0 & \frac{1+v_2 v_1}{\sqrt{(1-v_2^2)(1-v_1^2)}} \end{pmatrix} = B_{v_3}.$$

Hence we know that

$$\begin{cases} \frac{1+v_2 v_1}{\sqrt{(1-v_2^2)(1-v_1^2)}} = \frac{1}{\sqrt{1-v_3^2}}, \\ \frac{-v_2-v_1}{\sqrt{(1-v_2^2)(1-v_1^2)}} = \frac{-v_3}{\sqrt{1-v_3^2}}. \end{cases}$$

By dividing the second equality by the first equality, we get the result

$$v_3 = \frac{v_1 + v_2}{1 + v_1 v_2}.$$

7. Directly compute that $(B_v)^{-1} = B_{(-v)}$. If S' moves at a negative velocity v relative to S. Then we have S moves at a positive velocity v relative to S'. Let T_v be the transformation from S to S'. Then we have $[T_v^{-1}]_\beta = B_{(-v)}$ and so $[T_v]_\beta = B_v$.

8. In point of view of Earth, the astronaut should travel $2 \times 99/0.99 = 200$ years. So it will come back in the year 2200. However, let T_v and T_{-v} be the transformation from the space on Earth to the space of the astronaut

in the tour forward and the tour backward respectly. We may calculate that

$$
T_v \begin{pmatrix} 99 \\ 0 \\ 0 \\ 100 \end{pmatrix} = \frac{1}{\sqrt{1-0.99^2}} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 100 - 99 \times 0.99 \end{pmatrix} \cong \begin{pmatrix} 0 \\ 0 \\ 0 \\ 14.1 \end{pmatrix}.
$$

Hence the astronaut spent 14.1 years to travel to that star measured by himself. Similarly, we may compute that

$$
T_{-v} \begin{pmatrix} 99 \\ 0 \\ 0 \\ -100 \end{pmatrix} = \frac{1}{\sqrt{1-0.99^2}} \begin{pmatrix} 0 \\ 0 \\ 0 \\ -100 + 99 \times 0.99 \end{pmatrix} \cong \begin{pmatrix} 0 \\ 0 \\ 0 \\ -14.1 \end{pmatrix}.
$$

Hence he spent 14.1 years to travel back in his mind. Combining these two, he spent 28.2 years. Hence he would be return to Earth at age 48.2.

9. (a) The distance from earth to the star is b as measured on C.

   (b) At time t, the space-time coordinates of the star relative to S' and C' are

$$
B_v \begin{pmatrix} b \\ 0 \\ 0 \\ t \end{pmatrix} = \frac{1}{\sqrt{1-v^2}} \begin{pmatrix} b - vt \\ 0 \\ 0 \\ t - bv \end{pmatrix}.
$$

   (c) Compute $x' + t'v$ to eliminate the parameter t by

$$
x' - t'v = \frac{1}{\sqrt{1-v^2}} b(1 - v^2) = b\sqrt{1-v^2}.
$$

   Hence we get the result.

   (d) i. The speed that the star comes to the astronaut should be

$$
\frac{dx'}{dt'} = -v.
$$

   Hence the astronaut feel that he travels with the speed v.

   ii. In the astronaut's mind, he leave earth at $t' = 0$. Hence in S the earth is at $b\sqrt{1-v^2}$.

6.10   Conditioning and the Rayleigh Quotient

1. (a) No. The system $\begin{pmatrix} k^2 & 0 \\ 0 & 1 \end{pmatrix}$ is well-conditioned but its condition number is k, which can be arbitrarily large.

   (b) No. This is the contrapositive statement of the previous question.

(c) Yes. This is the result of Theorem 6.44.

(d) No. The norm of A is a value but the Rayleigh quotient is a function.

(e) No. See the Corollary 1 after Theorem 6.43. For example, the norm of $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ is 1 but the largest eigenvalue of it is 0.

2. Let A be the given matrix. Use the Corollary 1 to find the norm.

(a) The norm is $\sqrt{18}$.

(b) The norm is 6.

(c) The norm is $\dfrac{\sqrt{\sqrt{177}+15}}{\sqrt{6}}$.

3. If B is real symmetric, then we have $B^*B = B^2$. If λ is the largest eigenvalue of B, then we have λ^2 is also the largest eigenvalue of B^2. Apply the Corollary 1 after Theorem 6.43 and get that $\|B\| = \lambda$. If B is not real, the eigenvalue of B may not be a real number. So they are not comparable. Hence we need the condition that B is a real matrix here.

4. (a) Use the previous exercise. We know that $\|A\| = 84.74$ and $\|A^{-1}\| = \frac{1}{0.0588} \cong 17.01$. And the condition number is the ratio
$$\text{cond}(A) = \frac{84.74}{0.0588} \cong 1443.$$

(b) We have that
$$\|\tilde{x} - A^{-1}b\| = \|A^{-1}(b - A\tilde{x})\| \le \|A^{-1}\|\|b - A\tilde{x}\| = 17.01 \times 0.001 = 0.017$$
and
$$\frac{\|\tilde{x} - A^{-1}b\|}{\|A^{-1}b\|} \le \text{cond}(A)\frac{\|\delta b\|}{\|b\|} \cong \frac{14.43}{\|b\|}$$
by Theorem 6.44.

5. Use Theorem 6.44. Thus we get
$$\frac{\|\delta x\|}{\|x\|} \le \text{cond}(A)\frac{\|\delta b\|}{\|b\|} = 100 \times \frac{0.1}{1} = 10$$
and
$$\frac{\|\delta x\|}{\|x\|} \ge \frac{1}{\text{cond}(A)}\frac{\|\delta b\|}{\|b\|} = \frac{1}{100} \times \frac{0.1}{1} = 0.001.$$

6. Let $x = (1, -2, 3)$. First compute that
$$R(x) = \frac{\langle Bx, x \rangle}{\|x\|^2}$$
$$= \frac{\langle (3, 0, 5), (1, -2, 3) \rangle}{14} = \frac{9}{7}.$$

Second, compute the eigenvalues of B to be $4, 1, 1$. Hence we have $\|B\| = 4$ and $\|B^{-1}\| = 1^{-1} = 1$. Finally calculate the condition number to be

$$\mathrm{cond}(B) = 4 \times 1 = 4.$$

7. Follow the proof of Theorem 6.43. We have

$$R(x) = \frac{\sum_{i=1}^{n} \lambda_i |a_i|^2}{\|x\|^2} \geq \frac{\lambda_n \sum_{i=1}^{n} |a_i|^2}{\|x\|^2} = \frac{\lambda_n \|x\|^2}{\|x\|^2} = \lambda_n.$$

And the value is attainable since $R(v_n) = \lambda_n$.

8. Let λ be an eigenvalue of AA^*. If $\lambda = 0$, then we have AA^* is not invertible. Hence A and A^* are not invertible, so is A^*A. So λ is an eigenvalue of A^*A.

   Suppose now that $\lambda \neq 0$. We may find some eigenvector x such that $AA^*x = \lambda x$. This means that

   $$A^*A(A^*x) = \lambda A^*x.$$

   Since A^*x is not zero, λ is an eigenvalue of A^*A.

9. Since we have $A\delta x = \delta b$ and $x = A^{-1}b$, we have the inequalities

   $$\|\delta x\| \geq \frac{\|\delta b\|}{\|A\|}$$

   and

   $$\|x\| \leq \|A^{-1}\|\|b\|.$$

   Hence we get the inequality

   $$\frac{\|\delta x\|}{\|x\|} \geq \frac{\|\delta b\|}{\|A\|} \frac{1}{\|A^{-1}\|\|b\|} = \frac{1}{\|A\| \cdot \|A^{-1}\|} \frac{\|\delta b\|}{\|b\|}.$$

10. This is what we proved in the previous exercise.

11. If $A = kB$, then we have $A^*A = k^2 I$. So all the eigenvalues of A^*A are k^2. Thus we have $\|A\| = k$ and $\|A^{-1}\| = k^{-1}$. Hence the condition number of A is $k \cdot k^{-1} = 1$.

    Conversely, if $\mathrm{cond}(A) = 1$, we have $\lambda_1 = \lambda_n$ by Theorem 6.44. This means that all the eigenvalues of A^*A are the same. Denote the value of these eigenvalue by k. Since A^*A is self-adjoint, we could find an orthonormal basis $\beta = \{v_i\}$ consisting of eigenvectors. But this means that

    $$A^*A(v_i) = kv_i$$

    for all i. Since β is a basis, we get that actually $A^*A = kI$. This means that $B = \frac{1}{\sqrt{k}}A$ is unitary of orthogonal since $B^*B = I$. Thus A is a scalar multiple of B.

12. (a) If A and B are unitarily equivalent. We may write $B = Q^*AQ$ for some unitary matrix Q. Since Q is unitary, we have $\|Q(x)\| = \|x\|$. So we have
$$\frac{\|Bx\|}{\|x\|} = \frac{\|Q^*AQx\|}{\|x\|} = \frac{\|AQx\|}{\|Qx\|}.$$
Since any unitary matrix is invertible, we get the equality $\|A\| = \|B\|$.

(b) Write
$$\beta = \{v_1, v_2, \ldots, v_n\}$$
and
$$x = \sum_{i=1}^{n} a_i v_i.$$
We observe that
$$\|x\|^2 = \langle x, x \rangle = \sum_{i=1}^{n} a_i^2 = \|\phi_\beta(x)\|^2,$$
where $\phi_\beta(x)$ means the coordinates of x with respect to β. So we have
$$\|x\| = \|\phi_\beta(x)\|$$
This means that
$$\|T\| = \max_{x \neq 0} \frac{\|T(x)\|}{\|x\|} = \max_{x \neq 0} \frac{\|\phi_\beta(T(x))\|}{\|\phi_\beta(x)\|}$$
$$== \max_{\phi_\beta(x) \neq 0} \frac{\|[T]_\beta \phi_\beta(x)\|}{\|\phi_\beta(x)\|} = \|[T]_\beta\|.$$

(c) We have $\|T\| \geq k$ for all integer k since we have
$$\frac{\|T(v_k)\|}{\|v_k\|} = \frac{\|kv_k\|}{\|v_k\|} = k.$$

13. (a) If λ_1 is the largest eigenvalue of A^*A, then we know that $\sigma_i = \sqrt{\lambda_i} = \|A\|$.

(b) This comes from that the nonzero singular values of $A^\dagger$ are
$$\sigma_r^{-1} \geq \sigma_{r-1}^{-1} \geq \cdots \geq \sigma_1^{-1}.$$

(c) If A is invertible with the largest and the smallest eigenvalues of A^*A to be λ_1 and $\lambda_n > 0$, we know that $\sigma_1 = \sqrt{\lambda_1}$ and $\sigma_n = \sqrt{\lambda_n}$. Hence we have
$$\text{cond}(A) = \frac{\sigma_1}{\sigma_n}.$$

6.11 The Geometry of Orthogonal Operators

1. (a) No. It may be the compose of two or more rotations. For example, $T(x,y) = (R_\theta(x), R_\theta(y))$ is orthogonal but not a rotation or a reflection, where $x, y \in \mathbb{R}^2$ and R_θ is the rotation transformation about the angle θ.

    (b) Yes. See the Corollary after Theorem 6.45.

    (c) Yes. See Exercise 6.11.6.

    (d) No. For example, $U(x,y) = (R_\theta(x), y)$ and $T(x,y) = (x, R_\theta(y))$ are two rotations, where $x, y \in \mathbb{R}^2$ and R_θ is the rotation transformation about the angle θ. But UT is not a rotation.

    (e) Yes. It comes from the definition of a rotation.

    (f) No. In two-dimensional real space, the composite of two reflections is a rotation.

    (g) No. It may contains one reflection. For example, the mapping $T(x) = -x$ could not be the composite of rotations since the only rotation in $\mathbb{R}$ is the identity.

    (h) No. It may be the composite of some rotations and a reflection. For example, $T(x,y) = (-x, R_\theta(y))$ has $\det(T) = -1$ but it's not a reflection, where $x \in \mathbb{R}$, $y \in \mathbb{R}^2$, and R_θ is the rotation about the angle θ.

    (i) Yes. Let T be the reflection about W. We have that $W \oplus W^\perp = V$. So one of W and $W^\perp$ could not be zero. But every nonzero vector in W is an eigenvector with eigenvalue 1 while that in $W^\perp$ is an eigenvector with eigenvalue -1. So T must have eigenvalues.

    (j) No. The rotation on two-dimentional space has no eigenvector unless it's the identity mapping.

2. By Exercise 6.5.3 we know that the composite of two orthogonal operators should be an orthogonal operator.

3. (a) Check that $A^*A = AA^* = I$. So A is orthogonal. Hence it's a reflection by Theorem 6.45 since its determinant is -1.

    (b) Find the subspace $\{x : Ax = x\}$. That is, find the null space of $A - I$. Hence the axis is
    $$\text{span}\{(\sqrt{3}, 1)\}.$$

    (c) Compute $\det(B) = -1$. Hence we have $\det(AB) = \det(BA) = 1$. By Theorem 6.45, both of them are rotations.

4. (a) Compute that
    $$\det(A) = -\cos^2 \phi - \sin^2 \phi = -1.$$
    By Theorem 6.45, it's a reflection.

(b) Find the subspace $\{x : Ax = x\}$. That is, find the null space of $A - I$. Hence the axis is
$$\text{span}\{(\sin\phi, 1 - \cos\phi) =\}.$$

5. Let $\alpha = \{e_1, e_2\}$ be the standard basis in $\mathbb{R}^2$.

   (a) We may check that the rotation T_ϕ is a linear transformation. Hence it's enough to know
   $$\begin{cases} T(e_1) = (\cos\phi, \sin\phi), \\ T(e_2) = (-\sin\phi, \cos\phi) \end{cases}$$
   by directly rotate these two vectors. Hence we have $[T_\phi]_\alpha = A$.

   (b) Denote that
   $$A_\phi = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}.$$
   Directly compute that $A_\phi A_\psi = A_{\phi+\psi}$. So we have
   $$[T_\phi T_\psi]_\alpha = [T_\phi]_\alpha [T_\psi]_\alpha = A_\phi A_\psi = A_{\phi+\psi} = [T_{\phi+\psi}]_\alpha.$$

   (c) By the previous argument we kow that
   $$T_\phi T_\psi = T_{\phi+\psi} = T_{\psi+\phi} = T_\psi T_\phi.$$

6. If U and T are two rotations, we have $\det(UT) = \det(U)\det(T) = 1$. Hence by Theorem 6.47 UT contains no reflection. If V could be decomposed by three one-dimensional subspaces, they all of them are identities, thus UT is an identity mapping. Otherwise V must be decomposed into one one-dimensional and one two-dimensional subspaces. Thus UT is a rotation on the two-dimensional subspace and is an idenetiy on the one-dimensional space. Hence UT must be a rotation.

7. (a) We prove that if T is an orthogonal operator with $\det(T) = 1$ on a three-dimensional space V, then T is a rotation. First, we know that the decomposition of T contains no reflections by Theorem 6.47. According to T, V could be decomposed into some subspaces. If V is decomposed into three one-dimensional subspace, then T is the identity mapping on V since its an identity mapping on each subspace. Otherwise V should be decomposed into one one-dimensional and one two-dimensional subspaces. Thus T is a rotation on the two-dimensional subspace and is an idenetiy mapping on the one-dimensional space. Hence T must be a rotation.

   Finally, we found that $\det(A) = \det(B) = 1$. Hence they are rotations.

   (b) It comes from the fact that $\det(AB) = \det(A)\det(B) = 1$.

   (c) It should be the null space of $AB - I$,
   $$\text{span}\{((1 + \cos\phi)(1 - \cos\psi), (1 + \cos\phi)\sin\psi, \sin\phi\sin\psi)\}.$$

8. If T is an orthogonal operator, we know that the determinant of T should be ± 1. Now pick an orthonormal basis β. If $\det(T) = 1$, we have $\det([T]_\beta) = 1$ and hence $[T]_\beta$ is a rotation matrix by Theorem 6.23. By Exercise 6.2.15 we know that the mapping ϕ_{beta}, who maps $x \in V$ into its coordinates with respect to β, preserve the inner product. Hence T is a rotation when $[T]_\beta$ is a rotation. On the other hand, if $\det(T) = \det([T]) = -1$, we know that $[T]_\beta$ is a reflection matrix by Theorem 6.23. Again, T is a reflection since $[T]_\beta$ is a reflection matrix.

9. If T is a rotation, its decomposition contains no reflection. Hence we have $\det(T) = 1$ by Theorem 6.47. If T is a reflection, then we have its decomposition could contain exactly one reflection. Hence we have $\det(T) = -1$ by Theorem 6.47. So T cannot be both a rotation and a reflection.

10. If V is a two-dimensional real inner product space, we get the result by the Corollary after Theorem 6.45. If V is a three-dimensional real inner product space and U, T are two rotations, we have $\det(UT) = \det(U)\det(T) = 1$. By the discussion in Exercise 6.11.7(a), we know that UT should be a rotation.

11. Let $T(x, y) = (R_\theta(x), R_\theta(y))$ be an orthogonal operator, where $x, y \in \mathbb{R}^2$ and R_θ is the rotation transformation about the angle θ. It is neither a rotation nor a reflection.

12. Let β be an orthonormal basis. Then we have $[T]_\beta = -I_n$, where n is the dimsion of V. Since $\det(-I_n) = (-1)^n$, we know that T could decomposed into rotations if and only if n is even by the Corollary after Theorem 6.47.

13. Use the notation in that Lemma. We know that

$$W = \phi_\beta^{-1}(Z) = \text{span}\{\phi_\beta^{-1}(x_1), \phi_\beta^{-1}(x_2)\}.$$

And compute that

$$T(\phi_\beta^{-1}(x_i)) = \phi_\beta^{-1}(Ax_i) \in \phi_\beta^{-1}(Z)$$

for $i = 1, 2$. Hence W is T-invariant.

14. (a) It comes from that

$$\|T_W(x)\| = \|T(x)\| = \|x\|.$$

(b) Suppose y is an element in $W^\perp$. Since T_W is invertible by the previous argument, for each $x \in W$ we have $x = T(z)$ for some $z \in W$. This means that
$$\langle T(y), x \rangle = \langle y, T^*T(z) \rangle = \langle y, z \rangle = 0.$$

(c) It comes from that

$$\|T_{W^\perp}(x)\| = \|T(x)\| = \|x\|.$$

15. Let $t = 0$ in the equality in Theorem 5.24.

16. **(c)** As the definition of T_i given in the Corollary, we know that

$$T_i T_j(x) = \cdots + T(x_i) + \cdots + T(x_j) + \cdots T_j T_i(x),$$

where

$$x = x_1 + x_2 + \cdots + x_m.$$

(d) Again, we have

$$T(x) = T(x_1) + T(x_2) + \cdots + T(x_m) = T_1 T_2 \cdots T_m(x),$$

where

$$x = x_1 + x_2 + \cdots + x_m.$$

(e) We know that $\det(T_{W_i}) = \det(T_i)$ since $V = W_i \oplus W_i^\perp$ and

$$\det(T_{W_i^\perp}) = \det(I_{W_i^\perp}) = 1.$$

So T_i is a rotation if and only if T_{W_i} is a rotation. By Theorem 6.47 we get the result.

17. I think here we won't say an identity is a rotation. Otherwise the identity mapping could be decomposed into n identity mapping. Also, we need some fact. That is, if W_i is a subspace with dimension one in the decomposition, then T_{W_i} could not be a rotation since $T_{W_i}(x)$ sould be either x or $-x$. Hence every ration has the dimension of it subspace two.

   (a) By the Corollary after Theorem 6.47 we know that there is at most one reflection in the decomposition. To decompose a space with dimension n by rotations, there could be only $\frac{1}{2}(n-1)$ rotations.

   (b) Similarly, there is at most one reflection. If there's no reflection, then there're at most $\frac{1}{2}n$ rotations. If there's one reflection, there at most

$$\lfloor \frac{n-1}{2} \rfloor = \frac{1}{2}(n-2)$$

   rotations.

18. Let $\beta = \{x, x'\}$ be an orthonormal basis of V. Since $\|y\| = 1$, we may write $\phi_\beta(y) = (\cos\phi, \sin\phi)$ for some angle ϕ. Let

$$A_\phi = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

and T be the transformation with $[T]_\beta = A$. We have that $T(x) = y$ and T is a rotation.

On the other hand, by the definition of a rotation, we must have

$$T(x) = (\cos\theta)x + (\sin\theta)x'$$

and
$$T(x') = (-\sin\theta)x + (\cos\theta)x'.$$

Thus we must have $\cos\phi = \cos\theta$ and $\sin\phi = \sin\theta$. If $0 \le \phi, \theta < 2\pi$, we must have $\phi = \theta$. So the rotation is unique.

Chapter 7

Canonical Forms

7.1 The Jordan Canonical Form I

1. (a) Yes. It comes directly from the definition.

   (b) No. If x is a generalized eigenvector, we can find the smallest positive integer p such that $(T - \lambda I)^p(x) = 0$. Thus $y = (T - \lambda I)^{p-1} \neq 0$ is an eigenvector with respect to the eigenvalue λ. Hence λ must be an eigenvalue.

   (c) No. To apply the theorems in this section, the characteristic polynomial should split. For example, the matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ over $\mathbb{R}$ has no eigenvalues.

   (d) Yes. This is a result of Theorem 7.6.

   (e) No. The identity mapping I_2 from $\mathbb{C}^2$ to $\mathbb{C}^2$ has two cycles for the eigenvalue 1.

   (f) No. The basis β_i may not consisting of a union of cycles. For example, the transformation $T(a, b) = (a + b, b)$ has only one eigenvalue 1. The generalized eigenspace $K_1 = \mathbb{F}^2$. If

   $$\beta = \{(1, 1), (1, -1)\},$$

   then the matrix representation would be

   $$[T]_\beta = \frac{1}{2} \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix},$$

   which is not a Jordan form.

   (g) Yes. Let α be the standard basis. Then $[L_J]_\alpha = J$ is a Jordan form.

   (h) Yes. This is Theorem 7.2.

2. Compute the characteristic polynomial to find eigenvalues as what we did before. For each λ, find a basis for K_λ consisting of a union of disjoint cycles by computing bases for the null space of $(A - \lambda I)^p$ for each p. Write down the matrix S whose columns consist of these cycles of generalized eigenvectors. Then we will get the Jordan canonical form $J = S^{-1}AS$. When the matrix is diagonalizable, the Jordan canonical form should be the diagonal matrix sismilar to A. For more detail, please see the examples in the textbook. On the other hand, these results were computed by Wolfram Alpha. For example, the final question need the command below.

```
JordanDecomposition[{{2,1,0,0},{0,2,1,0},{0,0,3,0},{0,1,-1,3}}]
```

(a)
$$S = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}, J = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

(b)
$$S = \begin{pmatrix} -1 & 2 \\ 1 & 3 \end{pmatrix}, J = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}.$$

(c)
$$S = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 1 & 2 \\ 0 & 1 & 0 \end{pmatrix}, J = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

(d)
$$S = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{pmatrix}, J = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

3. Pick one basis β and write down the matrix representation $[T]_\beta$. Then do the same thing in the previous exercises. Again, we denote the Jordan canonical form by J and the matrix consisting of Jordan canonical basis by S. The Jordan canonical basis is the set of vector in V corresponding those column vectors of S in $\mathbb{F}^n$.

(a) Pick β to be the standard basis

$$\{1, x, x^2\}$$

and get

$$[T]_\beta = \begin{pmatrix} -2 & -1 & 0 \\ 0 & 2 & -2 \\ 0 & 0 & 2 \end{pmatrix}$$

and

$$S = \begin{pmatrix} 1 & -1 & \frac{1}{4} \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{pmatrix}, J = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

241

(b) Pick β to be the basis
$$\{1, t, t^2, e^t, te^t\}$$

and get
$$[T]_\beta = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

and
$$S = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, J = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

(c) Pick β to be the standard basis
$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

and get
$$[T]_\beta = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and
$$S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, J = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

(d) Pick β to be the standard basis
$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

and get
$$[T]_\beta = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

and
$$S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \end{pmatrix}, J = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

4. We may observe that $W = \text{span}(\gamma)$ is $(T - \lambda I)$-invariant by the definition of a cycle. Thus for all $w \in W$, we have

$$T(w) = (T - \lambda I)(w) + \lambda I(w) = (T - \lambda I)(w) + \lambda w \in W.$$

5. If x is an element of in two cycles, which is said to be γ_1 and γ_2 without lose of generality, we may find the smallest ingeter q such that

$$(T - \lambda I)^q(x) = 0.$$

This means that the initial eigenvectors of γ_1 and γ_2 are both $(T - \lambda I)^{q-1}(x)$. This is a contradiction. Hence all cycles are disjoint.

6.   (a) Use the fact that $T(x) = 0$ if only if $(-T)(x) = -0 = 0$.
  (b) Use the fact that $(-T)^k = (-1)^k T$.
  (c) It comes from the fact

$$(\lambda I_V - T)^k = [-(T - \lambda I_V)]^k$$

    and the previous argument.

7.   (a) If $U^k(x) = 0$, then $U^{k+1}(x) = U^k(U(x)) = 0$.
  (b) We know $U^{m+1}(V) = U^m(U(V)) \subset U^m(V)$. With the assumption $\text{rank}(U^m) = \text{rank}(U^{m+1})$, we know that

$$U^{m+1}(V) = U^m(V).$$

    This means

$$U(U^m(V)) = U^m(V)$$

    and so $U^k(V) = U^m(V)$ for all integer $k \geq m$.
  (c) The assumption $\text{rank}(U^m) = \text{rank}(U^{m+1})$ implies

$$\text{null}(U^m) = \text{null}(U^{m+1})$$

    by Dimension Theorem. This means $N(U^m) = N(U^{m+1})$ by the previous argument. If $U^{m+2}(x) = 0$, then $U(x)$ is an element in $N(U^{m+1}) = N(U^m)$. Hence we have $U^m(U(x)) = 0$ and thus x is an element in $N(U^{m+1})$. This means that $N(U^{m+2}) \subset N(U^{m+1})$ and so they are actually the same. Doing this inductively, we know that $N(U^m) = N(U^k)$ for all integer $k \geq m$.
  (d) By the definition of K_λ, we know

$$K_\lambda = \cup_{p \geq 1} N((T - \lambda I)^p).$$

    But by the previous argument we know that

$$N((T - \lambda I)^m) = N((T - \lambda I)^k)$$

    for all integer $k \geq m$ and the set is increasing as k increases. So actually K_λ is $N((T - \lambda I)^m)$.

(e) Since the characteristic polynomial splits, the transformation T is diagonalizable if and only if $K_\lambda = E_\lambda$. By the previous argument, we know that
$$K_\lambda = N(T - \lambda I) = E_\lambda.$$

(f) If λ is an eigenvalue of T_W, then λ is also an eigenvalue of T by Theorem 5.21. Since T is diagonalizable, we have the condition
$$\operatorname{rank}(T - \lambda I) = \operatorname{rank}((T - \lambda I)^2)$$

and so
$$N(T - \lambda I) = N((T - \lambda I)^2)$$

by the previous arguments. This implies why, which thereom used

$$N(T_W - \lambda I_W) = N((T_W - \lambda I_W)^2).$$

By Dimension Theorem, we get that
$$\operatorname{rank}(T_W - \lambda I_W) = \operatorname{rank}((T_W - \lambda I_W)^2).$$

So T_W is diagonalizable.

8. Theorem 7.4 implies that
$$V = \oplus_\lambda K_\lambda.$$

So the representation is uniques.

9. (a) The subspace W is T-invariant by Exercise 7.1.4. Let $\{v_i\}_i$ be the cycle with initial vector v_1. Then $[T_W]_\gamma$ is a Jordan block by the fact
$$T_W(v_1) = \lambda v_1$$

and
$$T_W(v_i) = v_{i-1} + \lambda v_i$$

for all $i > 1$. And β is a Jordan canonical basis since each cycle forms a block.

(b) If the ii-entry of $[T]_\beta$ is λ, then v_i is an nonzero element in K_λ. Since $K_\lambda \cap K_\mu = \{0\}$ for distinct eigenvalues λ and μ, we know that β' is exactly those v_i's such that the ii-entry of $[T]_\beta$ is λ. Let m be the number of the eigenvalue λ in the diagonal of $[T]_\beta$. Since a Jordan form is upper triangular. We know that m is the multiplicity of λ. By Theorem 7.4(c) we have
$$\dim(K_\lambda) = m = |\beta'|.$$

So β' is a basis for K_λ.

10. (a) Those initial vectors of q disjoint cycles forms a independent set consisting of eigenvectors in E_λ.

(b) Each block will correspond to one cycle. Use the previous argument and get the result.

11. By Theorem 7.7 and the assumptiong that the characteristic polynomial of L_A splits, the transformation L_A has a Jordan form J and a corresponding Jordan canonical basis β. Let α be the standarad basis for $\mathbb{F}^n$. Then we have
$$J = [L_A]_\beta = [I]_\alpha^\beta A [I]_\beta^\alpha$$
and so A is similar to J.

12. By Theorem 7.4(b), the space V is the direct sum of each $\mathrm{span}(\beta_i) = K_{\lambda_i}$, where β_i is a basis for K_{λ_i}.

13. Denote K_{λ_i} by W_i. Let β_i be a Jordan canonical basis for T_{W_i}. Apply Theorem 5.25 and get the result.

7.2   The Jordan Canonical Form II

1. (a) Yes. A diagonal matrix is a Jordan canonical form of itself. And the Corollary after Theorem 7.10 tolds us the Jordan canonical form is unique.

(b) Yes. This is a result of Theorem 7.11.

(c) No. The two matrices $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ have different Jordan canonical form. By Theorem 7.11, they cannot be similar.

(d) Yes. This is a result of Theorem 7.11.

(e) Yes. They are just two matrix representations of one transformation with different bases.

(f) No. The two matrices $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ have different Jordan canonical form.

(g) No. The identity mapping I from $\mathbb{C}^2$ to $\mathbb{C}^2$ has two different bases
$$\{(1,0), (0,1)\}$$
and
$$\{(1,1), (1,-1)\}.$$

(h) Yes. This is the Corollary after Theorem 7.10.

2. A coloumn in a dot diagram means a cycle. Each cycle has the corresponding eigenvalue as the diagonal entries and 1 as the upper subdiagonal

entries. So we have

$$J_1 = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix},$$

$$J_2 = \begin{pmatrix} 4 & 1 & 0 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix},$$

and

$$J_3 = \begin{pmatrix} -3 & 0 \\ 0 & -3 \end{pmatrix}.$$

And the Jordan canonical form J of T is

$$J = \begin{pmatrix} J_1 & O & O \\ O & J_2 & O \\ O & O & J_3 \end{pmatrix},$$

where O is the zero matrix with appropriate size.

3. (a) It's upper triangular. So the characteristic polynomial is $(2-t)^5(3-t)^2$.

   (b) Do the inverse of what we did in Ecercise 7.2.2. For $\lambda = 2$, the dot diagram is

   For $\lambda = 3$, the dot diagram is

   (c) The eigenvalue λ with it corresponding blocks are diagonal is has the property $E_\lambda = K_\lambda$.

   (d) The integer p_i is the length of the longest cycle in K_{λ_i}. So $p_2 = 3$ and $p_3 = 1$.

   (e) By Exercise 7.1.9(b), the matrix representations of U_2 and U_3 are

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

and

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

So we have $\text{rank}(U_2) = 3$, $\text{rank}(U_2^2) = 1$ and $\text{rank}(U_3) = \text{rank}(U_3^2) = 0$. By Dimension Theorem, we have $\text{null}(U_2) = 2$, $\text{null}(U_2^2) = 4$ and $\text{null}(U_3) = \text{null}(U_3^2) = 2$.

4. Do the same thing in Exercise 7.1.2.

   (a)
   $$Q = \begin{pmatrix} 1 & -1 & -1 \\ 2 & -1 & -2 \\ 1 & 1 & 0 \end{pmatrix}, J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

   (b)
   $$Q = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix}, J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

   (c)
   $$Q = \begin{pmatrix} 0 & -1 & \frac{2}{3} \\ -1 & -1 & -\frac{1}{3} \\ 1 & 3 & 0 \end{pmatrix}, J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

   (d)
   $$Q = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 1 & -1 & 0 & 1 \\ 1 & -2 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}, J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

5. For the following questions, pick some appropriate basis β and get the marix representation $A = [T]_\beta$. If A is a Jordan canonical form, then we've done. Otherwise do the same thing in the previous exercises. Similarly, we set $J = Q^{-1}AQ$ for some invertible matrix Q, where J is the Jordan canonical form. And the Jordan canonical basis is the set of vector in V corresponding those column vectors of Q in $\mathbb{F}^n$.

   (a) Pick the basis β to be

   $$\{e^t, te^t, \frac{1}{2}t^2e^t, e^{2t}\}$$

   and get the matrix representation

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

247

(b) Pick the basis β to be
$$\{1, x, \frac{x^2}{2}, \frac{x^3}{12}\}$$
and get the matrix representation
$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

(c) Pick the basis β to be
$$\{1, \frac{x^2}{2}, x, \frac{x^3}{6}\}$$
and get the matrix representation
$$A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

(d) Pick the basis β to be
$$\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\}$$
and get the matrix representation
$$A = \begin{pmatrix} 2 & 0 & 1 & 0 \\ 0 & 3 & -1 & 1 \\ 0 & -1 & 3 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$
Thus we have
$$Q = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & -2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{pmatrix}$$
and
$$J = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

(e) Pick the basis β to be
$$\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\}$$

248

and get the matrix representation

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

Thus we have

$$Q = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & -3 \\ 0 & 1 & 0 & 3 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

and

$$J = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix}.$$

(f) Pick the basis β to be

$$\{1, x, y, x^2, y^2, xy\}$$

and get the matrix representation

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

Thus we have

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

and

$$J = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

6. The fact $\text{rank}(M) = \text{rank}(M^t)$ for arbitrary square matrix M and the fact

$$(A^t - \lambda I)^r = [(A - \lambda I)^t]^r = [(A - \lambda I)^r]^t$$

tell us

$$\text{rank}((A - \lambda I)^r) = \text{rank}((A^t - \lambda I)^r).$$

By Theorem 7.9 we know that A and At have the same dot diagram and the same Jordan canonical form. So A and A^t are similar.

7. (a) Let γ' be the set $\{v_i\}_i = 1^m$. The desired result comes from the fact

$$T(v_i) = \lambda v_i + v_{i+1}$$

for $1 \leq i \leq m - 1$ and

$$T(v_m) = \lambda v_m.$$

(b) Let β be the standard basis for $\mathbb{F}^n$ and β' be the basis obtained from β by reversing the order of the vectors in each cycle in β. Then we have $[L_J]_\beta = J$ and $[L_J]_{\beta'} = J^t$. So J and J^t are similar.

(c) Since J is the Jordan canonical form of A, the two matrices J and A are similar. By the previous argument, J and J^t are similar. Finally, that A and J are similar implies that A^t and J^t are similar. Hence A and A^t are similar.

8. (a) Let β be the set $\{v_i\}_{i=1}^m$. Then we have the similar fact

$$T(cv_1) = \lambda cv_1$$

and

$$T(cv_i) = \lambda cv_i + cv_{i-1}.$$

So the matrix representation does not change and the new ordered basis is again a Jordan canonical basis for T.

(b) Since $(T - \lambda I)(y) = 0$, the vector $T(x + y) = T(x)$ does not change. Hence γ' is a cycle. And the new basis obtained from β by replacing γ by γ' is again a union of disjoint cycles. So it i sa Jordan canonical basis for T.

(c) Let $x = (-1, -1, -1, -1)$ and $y = (0, 1, 2, 0)$. Apply the previous argument and get a new Jordan canonical basis

$$\left\{ \begin{pmatrix} -1 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

9. (a) This is because we drawn the dot diagram in the order such that the length of cycles are decreasing.

(b) We know that p_j and r_i are decreasing as i and j become greater. So p_j is number of rows who contains more than or equal to j dots. Hence

$$p_j = \max\{i : r_i \geq j\}.$$

Similarly, r_i is the number of columns who contains more than or equal to i dots. Hence

$$r_i = \max\{j : p + j \geq i\}.$$

(c) It comes from the fact that p_j decreases.

(d) There is only one way to draw a diagram such that its i-th row contains exactly r_i dots. Once the diagram has been determined, those p_j's are determined.

10. (a) By Theorem 7.4(c), the dimension of K_λ is the multiplicity of λ. And the multiplicity of λ is the sum of the lengths of all the blocks corresponding to λ since a Jordan canonical form is always upper triangular.

(b) Since $E_\lambda \subset K_\lambda$, these two subspaces are the same if and only if the have the same dimension. The previous argument provide the desired result since the dimension of E_λ is the number of blocks corresponding to λ. The dimensions of them are the same if and only if all the related blocks have size 1×1.

11. It comes from the fact that

$$[T^p]_\beta = ([T]_\beta)^p.$$

12. Denote D_k to be the diagonal consisting of those ij-entries such that $i - j = k$. So D_0 is the original diagonal. If A is upper triangular matrix whose entries in D_0 are all zero, we have the fact that the entries of A^p in D_k, $0 \leq k < p$, are all zero. So A must be nilpotent.

13. (a) If $T^i(x) = 0$, then $T^{i+1}(x) = T^i(T(x)) = 0$.

(b) Pick β_1 to be one arbitrary basis for $N(T^1)$. Extend β_i to be a basis for $N(T^{i+1})$. Doing this inductively, we get the described sequence.

(c) By Exercise 7.1.7(c), we know $N(T^i) \neq N(T^{i+1})$ for $i \leq p - 1$. And the desired result comes from the fact that

$$T(\beta_i) \subset N(T^{i-1}) = \text{span}(\beta_{i-1}) \neq \text{span}(\beta_i).$$

(d) The form of the characteristic polynomial directly comes from the previous argument. And the other observation is natural if the characteristic polynomial has been fixed.

14. Since the characteristic polynomial of T splits and contains the unique zero to be 0, T has the Jordan canonical form J whose diagonal entries are all zero. By Exercise 7.2.12, the matrix J is nilpotent. By Exercise 7.2.11, the linear operator T is also nilpotent.

15. The matrix
$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$
has the characteristic polynomial to be $-t(t^2 + 1)$. Zero is the only eigenvalue of $T = L_A$. But T and A is not nilpotent since $A^3 = -A$. By Exercise 7.2.13 and Exercise 7.2.14, a linear operator T is not nilpotent if and only if the characteristic polynomial of T is not of the form $(-1)^n t^n$.

16. Since the eigenvalue is zero now, observe that if x is an element in β corresponding to one dot called p, then $T(x)$ would be the element corresponding to the dot just above the dot p. So the set described in the exercise is an independent set in $R(T^i)$. By counting the dimension of $R(T^i)$ by Theorem 7.9, we know the set is a basis for $R(T^i)$.

17. (a) Assume that
$$x = v_1 + v_2 + \cdots + v_k$$
and
$$y = u_1 + u_2 + \cdots + u_k.$$
Then S is a linear operator since
$$S(x + cy) = \lambda_1(v_1 + cu_1) + \cdots + \lambda_k(v_k + cu_k)$$
$$= S(x) + cS(y).$$
Observe that if v is an element in K_λ, then $S(v) = \lambda v$. This means that if we pick a Jordan canonical basis β of T for V, then $[S]_\beta$ is diagonal.

(b) Let β be a Jordan canonical basis for T. By the previous argument we have $[T]_\beta = J$ and $[S]_\beta = D$, where J is the Jordan canonical form of T and D is the diagonal matrix given by S. Also, by the definition of S, we know that $[U]_\beta = J - D$ is an upper triangular matrix with each diagonal entry equal to zero. By Exercise 7.2.11 and Exercise 7.2.12 the operator U is nilpotent. And the fact U and S commutes is due to the fact $J - D$ and D commutes. The later fact comes from some direct computation.

18. Actually, this exercise could be a lemma for Exercise 7.2.17.

(a) It is nilpotent by Exercise 7.2.12 since M is a upper triangular matrix with each diagonal entry equal to zero.

(b) It comes from some direct computation.

(c) Since $MD = DM$, we have

$$J^r = (M + D)^r = \sum_{k=0}^{r} \binom{r}{k} M^k D^{r-k}.$$

The second equality is due to the fact that $M^k = O$ for all $k \geq p$.

19. (a) It comes from some direct computation. Multiplying N at right means moving all the columns to their right columns.

(b) Use Exercise 7.2.18(c). Now the matrix M is the matrix N in Exercise 7.2.19(a).

(c) If $|\lambda| < 1$, then the limit tends to a zero matrix. If $\lambda = 1$ amd $m = 1$, then the limit tends to the identity matrix of dimension 1. Conversely, if $|\lambda| \geq 1$ but $\lambda \neq 1$, then the diagonal entries will not converge. If $\lambda = 1$ but $m > 1$, the 12-entry will diverge.

(d) Observe the fact that if J is a Jordan form consisting of several Jordan blocks J_i, then $J^r = \oplus_i J_i^r$. So the $\lim_{m \to \infty} J^m$ exsist if and only if $\lim_{m \to \infty} J_i^m$ exists for all i. On the other hand, if A is a square matrix with complex entries, then it has the Jordan canonical form $J = Q^{-1}AQ$ for some Q. This means that $\lim_{m \to \infty} A^m$ exists if and only if $\lim_{m \to \infty} J^m$ exists. So Theorem 5.13 now comes from the result in Exercise 7.2.19(c).

20. (a) The norm $\|A\| \geq 0$ since $|A_{ij}| \geq 0$ for all i and j. The norm $\|A\| = 0$ if and only if $|A_{ij}| = 0$ for all i and j. So $\|A\| = 0$ if and only if $A = O$.

(b) Compute

$$\|cA\| = \max\{|cA_{ij}|\} = |c| \max\{|A_{ij}|\} = |c|\|A\|.$$

(c) It comes from the fact

$$|A_{ij} + B_{ij}| \leq |A_{ij}| + |B_{ij}|$$

for all i and j.

(d) Compute

$$\|AB\| = \max\{|(AB)_{ij}|\} = \max\{|\sum_{k=1}^{n} A_{ik}B_{kj}|\}$$

$$\leq \max\{|\sum_{k=1}^{n} \|A\|\|B\|\|\} = n\|A\|\|B\|.$$

21. (a) The Corollary after 5.15 implies that A^m is again a transition matrix. So all its entris are no greater than 1.

(b) Use the inequaliy in Exercise 7.2.20(d) and the previous argument. We compute

$$\|J^m\| = \|P^{-1}A^mP\| \le n^2\|P^{-1}\|\|A^m\|\|P\| \le n^2\|P^{-1}\|\|P\|.$$

Pick the fixed value $c = n^2\|P^{-1}\|\|P\|$ and get the result.

(c) By the previous argument, we know the norm $\|J^m\|$ is bounded. If J_1 is a block corresponding to the eigenvalue 1 and the size of J_1 is greater than 1, then the 12-entry of J_1^m is unbounded. This is a contradiction.

(d) By the Corollary 3 after Theorem 5.16, the absolute value of eigenvalues of A is no greater than 1. So by Theorem 5.13, the limit $\lim_{m \to \infty} A^m$ exists if and only if 1 is the only eigenvalue of A.

(e) Theorem 5.19 confirm that $\dim(E_1) = 1$. And Exercise 7.2.21(c) implies that $K_1 = E_1$. So the multiplicity of the eigenvalue 1 is equal to $\dim(K_1) = \dim(E_1) = 1$ by Theorem Theorem 7.4(c).

22. Since A is a matrix over complex numbers, A has the Jordan canonical form $J = Q^{-1}AQ$ for some invertible matrix Q. So e^A exists if e^J exist. Observe that

$$\|J^m\| \le n^{m-1}\|J\|$$

by Exercise 7.20(d). This means the sequence in the definition of e^J converge absolutely. Hence e^J exists.

23. (a) For brevity, we just write λ instead of λ_i. Also denote u_k to be $(A - \lambda I)^k u$. So we have $(A - \lambda I)u_k = u_{k+1}$. Let

$$x = e^{\lambda t}\Big[\sum_{k=0}^{p-1} f^{(k)}(t)u_{p-1-k}\Big]$$

be the function vector given in this question. Observe that

$$(A - \lambda I)x = e^{\lambda t}\Big[\sum_{k=0}^{p-1} f^{(k)}(t)u_{p-k}\Big].$$

Then compute that

$$x' = \lambda x + e^{\lambda t}\Big[\sum_{k=0}^{p-1} f^{(k+1)}(t)u_{p-1-k}\Big]$$

$$= \lambda x + e^{\lambda t}\Big[\sum_{k=1}^{p-1} f^{(k)}(t)u_{p-k}\Big]$$

$$= \lambda x + (A - \lambda I)x = Ax.$$

254

(b) Since A is a matrix over $\mathbb{C}$, it has the Jordan canonical form $J = Q^{-1}AQ$ for some invertible matrix Q. Now the system become

$$x' = QJQ^{-1}x$$

and so

$$Q^{-1}x' = JQ^{-1}x.$$

Let $y = Q^{-1}x$ and so $x = Qy$. Rewrite the system to be

$$y' = Jy.$$

Since the solution of y is the linear combination of the solutions of each Jordan block, we may just assume that J consists only one block. Thus we may solve the system one by one from the last coordinate of y to the first coordinate and get

$$y = e^{\lambda t}\begin{pmatrix} f(t) \\ f^{(1)}(t) \\ \vdots \\ f^{(p-1)}(t) \end{pmatrix}.$$

On the other hand, the last column of Q is the end vector u of the cycle. Use the notation in the previous question, we know Q has the form

$$Q = \begin{pmatrix} | & | & \cdots & | \\ u_{p-1} & u_{p-2} & \cdots & u_0 \\ | & | & \cdots & | \end{pmatrix}.$$

Thus the solution must be $x = Qy$. And the solution coincide the solution given in the previous question. So the general solution is the sum of the solutions given by each end vector u in different cycles.

24. As the previous exercise, we write $Q^{-1}AQ = J$, where J is the Jordan canonical form of A. Then solve $Y' = JY$. Finally the answer should be $X = QY$.

(a) The coefficient matrix is

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

Compute

$$J = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

and

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}.$$

255

Thus we know that

$$Y = e^{2t} \begin{pmatrix} at + b \\ a \\ 0 \end{pmatrix} + e^{3t} \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix}.$$

and so the solution is

$$X = QY.$$

(b) The coefficient matrix is

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

So $J = A$ and $Q = I$. Thus we know that

$$Y = e^{2t} \begin{pmatrix} at^2 + bt + c \\ 2at + b \\ 2a \end{pmatrix}$$

and so the solution is

$$X = QY.$$

7.3   The Minimal Polynomial

1. (a) No. If $p(t)$ is the polynomial of largest degree such that $p(T) = T_0$, then $q(t) = t(p(t)$ is a polynomial of larger degree with the same property $q(T) = T_0$.

   (b) Yes. This is Theorem 7.12.

   (c) No. The minimal polynomial divides the characteristic polynomial by Theorem 7.12. For example, the identity transformation I from $\mathbb{R}^2$ to $\mathbb{R}^2$ has its *characteristicpolynomial* $(1 - t)^2$ and its minimal polynomial $t - 1$.

   (d) No. The identity transformation I from $\mathbb{R}^2$ to $\mathbb{R}^2$ has its *characteristicpolynomial* $(1 - t)^2$ but its minimal polynomial $t - 1$.

   (e) Yes. Since f splits, it consists of those factors $(t - \lambda)^r$ for some $r \leq n$ and for some eigenvalues λ. By Theorem 7.13, the minimal polynomial p also contains these factors. So f divides p^n.

   (f) No. For example, the identity transformation I from $\mathbb{R}^2$ to $\mathbb{R}^2$ has its *characteristicpolynomial* $(1 - t)^2$ and its minimal polynomial $t - 1$.

   (g) No. For the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, its minimal polynomial is $(t - 1)^2$ but it is not diagonalizable.

   (h) Yes. This is Theorem 7.15.

(i) Yes. By Theorem 7.14, the minimal polynomial contains at least
n zeroes. Hence the degree of the minimal polynomial of T must
be greater than or equal to n. Also, by Cayley-Hamilton Theorem,
the degree is no greater than n. Hence the degree of the minimal
polynomial of T must be n.

2. Let A be the given matrix. Find the eigenvalues λ_i of A. Then the
minimal polynomial should be of the form $\prod_i (t - \lambda_i)_i^r$. Try all possible
r_i's. Another way is to compute the Jordan canonical form.

(a) The eigenvalues are 1 and 3. So the minimal polynomial must be
$(t-1)(t-3)$.

(b) The eigenvalues are 1 and 1. So the minimal polynomial could be
$(t-1)$ or $(t-1)^2$. Since $A - I \neq O$, the minimal polynomial must be
$(t-1)^2$.

(c) The Jordan canonical form is

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

So the minimal polynomial is $(t-2)(t-1)^2$.

(d) The Jordan canonical form is

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

So the minimal polynomial is $(t-2)^2$.

3. Write down the matrix and do the same as that in the previous exercise.

(a) The minimal polynomial is $(t - \sqrt{2})(t + \sqrt{2})$.

(b) The minimal polynomial is $(t-2)^3$.

(c) The minimal polynomial is $(t-2)^2$.

(d) The minimal polynomial is $(t+1)(t-1)$.

4. Use Theorem 7.16. So those matrices in Exercises 7.3.2(a), 7.3.3(a), and
7.3.3(d) are diagonalizable.

5. Let $f(t) = t^3 - 2t + t = t(t-1)^2$. Thus we have $f(T) = T_0$. So the
minimal polynomial $p(t)$ must divide the polynomial $f(t)$. Since T is
diagonalizable, $p(t)$ could only be t, $(t-1)$, or $t(t-1)$. If $p(t) = t$, then
$T = T_0$. If $p(t) = (t-1)$, then $T = I$. If $p(t) = t(t-1)$, then $[T]_\beta = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$
for some basis β.

6. Those results comes from the fact $[f(T)]_\beta = O$ if and only if $f(T) = T_0$.

7. By Theorem 7.12, $p(t)$ must of that form for some $m_i \le n_i$. Also, by Theorem 7.14, we must have $m_i \ge 1$.

8. (a) Let $f(t)$ be the characteristic polynomial of T. Recall that $\det(T) = f(0)$ by Theorem 4.4.7. So By Theorem 7.12 and Theorem 7.14 0 is a zero for $p(t)$ if and only if 0 is a zero for $f(t)$. Thus T is invertible if and only if $p(0) \ne 0$.

   (b) Directly compute that

   $$-\frac{1}{a_0}(T^{n-1} + a_{n-1}T^{n-2} + \cdots + a_2T + a_1I)T$$

   $$= -\frac{1}{a_0}(p(T) - a_0) = I.$$

9. Use Theorem 7.13. We know that V is a T-cyclic subspace if and only if the minimal polynomial $p(t) = (-1)^n f(t)$, where n is the dimension of V and f is the characteristic polynomial of T. Assume the characteristic polynomial $f(t)$ is

   $$(t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2}\cdots(t - \lambda_k)^{n_k},$$

   where n_i is the dimension of the eigenspace of λ_i since T is diagonalizable. Then the minimal polynomial must be

   $$(t - \lambda_1)(t - \lambda_2)\cdots(t - \lambda_k).$$

   So V is a T-cyclic subspace if and only if $n_i = 1$ for all i.

10. Let $p(t)$ be the minimal polynomial of T. Thus we have

    $$p(T_W)(w) = p(T)(w) = 0$$

    for all $w \in W$. This means that $p(T_W)$ is a zero mapping. Hence the minimal polynomial of T_W divides $p(t)$.

11. (a) If $y \in V$ is a solution to the equation $g(D)(y) = 0$, then $g(D)(y') = (g(D)(y))' = 0 \in V$.

    (b) We already know that
    $$g(D)(y) = 0$$

    for all $y \in V$. So the minimal polynomial $p(t)$ must divide $g(t)$. If the degree of $p(t)$ is less than but not equal to the degree of $g(t)$, then the solution space of the equation $p(D)(y) = 0$ must contain V. This will make the dimension of the solution space of $p(D)(y) = 0$ greater than the degree of $p(t)$. This is a contradiction to Theorem 2.32. Hence we must have $p(t) = g(t)$.

258

(c) By Theorem 2.32 the dimension of V is n, the degree of $g(t)$. So by Theorem 7.12, the characteristic polynomial must be $(-1)^n g(t)$.

12. Suppose, by contradiction, there is a polynomial $g(t)$ of degree n such $g(D) = T_0$. Then we know that $g(D)(x^n)$ is a constant but not zero. This is a contradiction to the fact $g(D)(x^n) = T_0(x^n) = 0$. So D has no minimal polynomial.

13. Let $p(t)$ be the polynomial given in the question. And let β be a Jordan basis for T. We have $(T - \lambda_i)^{p_i}(v) = 0$ if v is a generalized eigenvector with respect to the eigenvalue λ_i. So $p(T)(\beta) = \{0\}$. Hence the minimal polynomial $q(t)$ of T must divide $p(t)$ and must be of the form

$$(t - \lambda_1)^{r_1}(t - \lambda_2)^{r_2}\cdots(t - \lambda_k)^{r_k},$$

where $1 \le r_i \le p_i$. If $r_i < p_i$ for some i, pick the end vector u of the cycle of length p_i in β corresponding to the eigenvalue λ_i. This u exist by the definition of p_i. Thus $(T - \lambda_i)^{r_i}(u) = w \ne 0$. Since K_{λ_i} is $(T - \lambda_j)$-invariant and $T - \lambda_j$ is injective on K_{λ_i} for all $j \ne i$ by Theorem 7.1, we know that $q(T)(u) \ne 0$. Hence r_i must be p_i. And so $p(t)$ must be the minimal polynomial of T.

14. The answer is no. Let T be the identity mapping on $\mathbb{R}^2$. And let W_1 be the x-axis and W_2 be the y-axis. The minimal polynomial of T_{W_1} and T_{W_2} are both $t - 1$. But the minimal polynomial of T is $(t - 1)$ but not $(t - 1)^2$.

15. (a) Let W be the T-cyclic subspace generated by x. And let $p(t)$ be the minimal polynomial of T_W. We know that $p(T_W) = T_0$. If $q(t)$ is a polynomial such that $q(T)(x) = 0$, we know that $q(T)(T^k(x)) = T^k(q(T)(x)) = 0$ for all k. So $p(t)$ must divide $q(t)$ by Theorem 7.12. Hence $p(t)$ is the unique T-annihilator of x.

(b) The T-annihilator of x is the minimal polynomial of T_W by the previous argument. Hence it divides the characteristic polynomial of T, who divides any polynomial for which $g(T) = T_0$, by Theorem 7.12.

(c) This comes from the proof in Exercise 7.3.15(c).

(d) By the result in the previous question, the dimension of the T-cyclic subspace generated by x is equal to the degree of the T-annihilator of x. If the dimension of the T-cyclic subspace generated by x has dimension 1, then $T(x)$ must be a multiple of x. Hence x is an eigenvector. Conversely, if x is an eigenvector, then $T(x) = \lambda x$ for some λ. This means the dimension of the T-cyclic subspace generated by x is 1.

16. (a) Let $f(t)$ be the characteristic polynomial of T. Then we have $f(T)(x) = T_0(x) = 0 \in W_1$. So there must be some monic polynomial $p(t)$ of least positive degree for which $p(T)(x) \in W_1$. If $h(t)$ is a polynomial for which $h(T)(x) \in W_1$, we have

$$h(t) = p(t)q(t) + r(t)$$

259

for some polynomial $q(t)$ anr $r(t)$ such that the degree of $r(t)$ is less than the degree of $p(t)$ by Division Algorithm. This means that

$$r(T)(x) = h(T)(x) - p(T)p(T)(x) \in W_1$$

since W_1 is T-invariant. Hence the degree of $r(t)$ must be 0. So $p(t)$ divides $h(t)$. Thus $g_1(t) = p(t)$ is the unique monice polynomial of least positive degree such that $g_1(T)(x) \in W_1$.

(b) This has been proven in the previous argument.

(c) Let $p(t)$ and $f(t)$ be the minimal and characteristic polynomials of T. Then we have $p(T)(x) = f(T)(x) = 0 \in W_1$. By the previous question, we get the desired conclusion.

(d) Observe that $g_2(T)(x) \in W_2 \subset W_1$. So $g_1(t)$ divides $g_2(t)$ by the previous arguments.

7.4 The Rational Canonical Form

1. (a) Yes. See Theorem 7.17.

(b) No. Let $T(a,b) = (a, 2b)$ be a transformation. Then the basis $\beta = \{(1,1), (1,2)\}$ is a T-cyclic basis generated by $(1,1)$. But it is not a rational canonical basis.

(c) No. See Theorem 7.22.

(d) Yes. If A is a square matrix with its rational canonical form C with rational canonical basis β, then we have $C = [L_A]_\beta$.

(e) Yes. See Theorem 7.23(a).

(f) No. They are in general diffrent. For example, the dot diagram of the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has only one dot. It could not forms a basis.

(g) Yes. The matrix is similar to its Jordan canonical form and its rational canonical form. Hence the two forms should be similar.

2. Find the factorization of the characteristic polynomial. Find the basis of K_ϕ for each some monic irreducibla polyomial factor consisting of T-cyclic bases through the proof of Theorem 7.22. Write down the basis with some appropriate order as the columns of Q. Then compute $C = Q^{-1}AQ$ or find C by the dot diagram. And I want to emphasize that I compute these answers in Exercises 7.4.2 and 7.4.3 by HAND!

(a) It is a Jordan canonical form. So

$$Q = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 6 \\ 1 & 3 & 9 \end{pmatrix}$$

and
$$C = \begin{pmatrix} 0 & 0 & 27 \\ 1 & 0 & -27 \\ 0 & 1 & 9 \end{pmatrix}.$$

(b) It has been already the rational canonical form since the character-istic polynomial $t^2 + t + 1$ is irreducible in $\mathbb{R}$. So $C = A$ and $Q = I$.

(c) It is diagonalizable in $\mathbb{C}$. So
$$Q = \begin{pmatrix} 1 & 1 \\ \frac{\sqrt{3}\,i+1}{2} & \frac{1-\sqrt{3}\,i}{2} \end{pmatrix}$$
and
$$C = \begin{pmatrix} -\frac{\sqrt{3}\,i+1}{2} & 0 \\ 0 & \frac{\sqrt{3}\,i-1}{2} \end{pmatrix}.$$

(d) Try the generating vector $(1,0,0,0)$. So
$$Q = \begin{pmatrix} 1 & 0 & -7 & -4 \\ 0 & 1 & -4 & -3 \\ 0 & 0 & -4 & -4 \\ 0 & 0 & -4 & -8 \end{pmatrix}$$
and
$$C = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

(e) Use $(0,-1,0,1)$ and $(3,1,1,0)$ as generating vectors. So
$$Q = \begin{pmatrix} 0 & -3 & 3 & 8 \\ -1 & -2 & 1 & 5 \\ 0 & -3 & 1 & 5 \\ 1 & -4 & 0 & 7 \end{pmatrix}$$
and
$$C = \begin{pmatrix} 0 & -2 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

3. Write down the matrix representation A by some basis β and find the rational $C = Q^{-1}AQ$ for some inveritble Q. Then the rational canonical basis is the basis corresponding the columns of Q.

(a) Let
$$\beta = \{1, x, x^2, x^3\}.$$

Then
$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & -1 & -2 \end{pmatrix}$$

and
$$C = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

(b) Let $\beta = S$. Then
$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & -1 & -2 \end{pmatrix}$$

and
$$C = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

(c) Let
$$\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

Then
$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

and
$$C = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

(d) Let
$$\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

Then
$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

and

$$C = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

(e) Let $\beta = S$. Then

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

and

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

4. (a) We may write an element in $R(\phi(T))$ as $\phi(T)(x)$ for some x. Since $(\phi(T))^m(v) = 0$ for all v, we have $(\phi(T))^{m-1}(\phi(T)(x)) = (\phi(T))^m(x) = 0$.

(b) The matrix

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

has minimal polynomial $(t-1)^2$. Compute $R(L_A-I) = \mathrm{span}\{(1,0,0)\}$. But $(0,0,1)$ is an element in $N(L_A - I)$ but not in $R(L_A - I)$.

(c) We know that the minimal polynomial $p(t)$ of the restriction of T divides $(\phi(t))^m$ by Exercise 7.3.10. Pick an element x such that $(\phi(T))^{m-1}(x) \neq 0$. Then we know that $y = \phi(T)(x)$ is an element in $R(\phi(T))$ and $(\phi(T))^{m-2}(y) \neq 0$. Hence $p(t)$ must be $(\phi(t))^{m-1}$.

5. If the rational canonical form of T is a diagonal matrix, then T is diagonalizable naturally. Conversely, if T is diagonalizable, then the characteristic polynomial of T splits and $E_\lambda = K_{\phi_\lambda}$, where $\phi_\lambda = t - \lambda$, for each eigenvalue λ. This means each cyclic basis in K_{ϕ_λ} is of size 1. That is, a rational canonical basis consisting of eigenvectors. So the rational canonical form of T is a diagonal matrix.

6. Here we denote the degree of ϕ_1 and ϕ_2 by a and b respectly.

(a) By Theorem 7.23(b) we know the dimension of K_{ϕ_1} and K_{ϕ_2} are a and b respectly. Pick a nonzero element v_1 in K_{ϕ_1}. The T-annihilator of v_1 divides ϕ_1. Hence the T-annihilator of v_1 is ϕ_1. Find the nonzero vector v_2 in K_{ϕ_2} similarly such that the T-annihilator of v_2 is ϕ_2. Thus $\beta_{v_1} \cup \beta_{v_2}$ is a basis of V by Theorem 7.19 and the fact that $|\beta_{v_1} \cup \beta_{v_2}| = a + b = n$.

(b) Pick $v_3 = v_1 + v_2$, where v_1 and v_2 are the two vectors given in the previous question. Since $\phi_1(T)(v_2) \neq 0$ and $\phi_2(T)(v_1) \neq 0$ by Theorem 7.18. The T-annihilator of v_3 cannot be ϕ_1 and ϕ_2. So the final possibility of the T-annihilator is $\phi_1\phi_2$.

(c) The first one has two blocks but the second one has only one block.

7. By the definition of m_i, we know the

$$N(\phi_i(T)^{m_i-1} \neq N(\phi_i(T)^{m_i} = N(\phi_i(T)^{m_i+1} = K_{\phi_i}.$$

Apply Theorem 7.24 and get the result.

8. If $\phi(T)$ is not injective, we can find a nonzero element x such that $\phi(T)(x) = 0$. Hence the T-annihilator $p(t)$ divides $\phi(t)$ by Exercise 7.3.15(b). This means $p(t) = \phi(t)$. If $f(t)$ is the characteristic polynomial of T, then we have $f(T)(x) = 0$ by Cayley-Hamilton Theorem. Again by Exercise 7.3.15(b) we have $\phi(t)$ divides $f(t)$.

9. Since the disjoint union of β_i's is a basis, each β_i is independent and forms a basis of $\text{span}(\beta_i)$. Now denote

$$W_i = \text{span}(\gamma_i) = \text{span}(\beta_i).$$

Thus $V = \oplus_i W_i$ by Theorem 5.10. And the set $\gamma = \cup_i \gamma_i$ is a basis by Exercise 1.6.33.

10. Since $x \in C_y$, we may assume $x = T^m(y)$ for some integer m. If $\phi(t)$ is the T-annihilator of x and $p(t)$ is the T-annihilator of y, then we know $p(T)(x) = p(T)(T^m(y)) = T^m(p(T)(y)) = 0$. Hence $p(t)$ is a factor of $\phi(t)$. If $x = 0$, then we have $p(t) = 1$ and $y = 0$. The statement is true for this case. So we assume that $x \neq 0$. Thus we know that $y \neq 0$ otherwise x is zero. Hence we know $p(t) = \phi(t)$. By Exercise 7.3.15, the dimension of C_x is equal to the dimension of C_y. Since $x = T^m(y)$ we know that $C_x \subset C_y$. Finally we know that they are the same since they have the same dimension.

11. (a) Since the rational canonical basis exists, we get the result by Theorem 5.10.

(b) This comes from Theorem 5.25.

12. Let $\beta = \cup_i C_{v_i}$. The statement holds by Theorem 5.10.

GNU Free Documentation License

Preamble

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "**Document**", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "**you**". You accept the

license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "**Modified Version**" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "**Secondary Section**" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "**Invariant Sections**" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "**Cover Texts**" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "**Transparent**" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "**Opaque**".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "**Title Page**" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title

page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

The "**publisher**" means any person or entity that distributes copies of the Document to the public.

A section "**Entitled XYZ**" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "**Acknowledgements**", "**Dedications**", "**Endorsements**", or "**History**".) To "**Preserve the Title**" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.

B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.

C. State on the Title page the name of the publisher of the Modified Version, as the publisher.

D. Preserve all the copyright notices of the Document.

E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.

F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.

G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.

H. Include an unaltered copy of this License.

I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.

M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.

N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.

O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties—for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover

Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, or distribute it is void, and will automatically terminate your rights under this License.

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, receipt of a copy of some or all of the same material does not give you any rights to use it.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy's public statement of acceptance of a version permanently authorizes you to choose that version for the Document.

11. RELICENSING

"Massive Multiauthor Collaboration Site" (or "MMC Site") means any World Wide Web server that publishes copyrightable works and also provides prominent facilities for anybody to edit those works. A public wiki that anybody can edit is an example of such a server. A "Massive Multiauthor Collaboration" (or "MMC") contained in the site means any set of copyrightable works thus published on the MMC site.

"CC-BY-SA" means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a not-for-profit corporation with a principal place of business in San Francisco, California, as well as future copyleft versions of that license published by that same organization.

"Incorporate" means to publish or republish a Document, in whole or in part, as part of another Document.

An MMC is "eligible for relicensing" if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC, and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing.

ADDENDUM: How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

> Copyright © YEAR YOUR NAME. Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled "GNU Free Documentation License".

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the "with . . . Texts." line with this:

> with the Invariant Sections being LIST THEIR TITLES, with the Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.

Appendices

Some Definition in Graph Theory

Definition. A **graph** is a ordered pair of sets (V, E) and E is the set of pair of elements of V. A **directed graph**, or a **digraph** is a ordered pair of sets (V, E) and E is the set of ordered pair of elements of V. And we call V be the set of vertices and E be the set of edges. A **tournament** is a digraph such that for each pair of vertices we have exactly one edge connecting them.

Example. Let $G(V_1, E_1)$ be a graph with vertex set $V_1 = \{1, 2, 3, 4\}$ and edge set
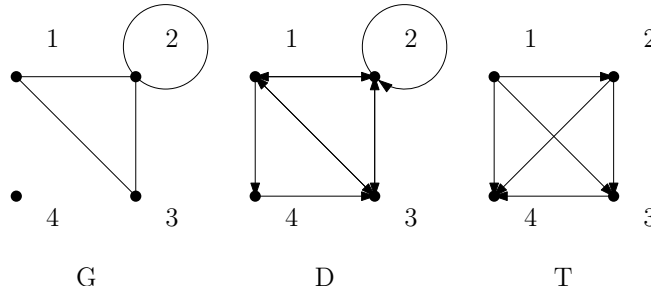$$E_1 = \{(1, 2), (2, 2), (2, 3), (3, 1)\}.$$
And let $D(V_2, E_2)$ be a *digraph* with vertex set $V_2 = \{1, 2, 3, 4\}$ and edge set
$$E_2 = \{(1, 2), (2, 1), (2, 2), (2, 3), (3, 2), (1, 3), (3, 1), (1, 4), (4, 3)\}.$$
Finally let $T(V_3, E_3)$ be a digraph with vertex set $V_2 = \{1, 2, 3, 4\}$ and edge set
$$E_3 = \{(1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4)\}.$$
Thus T is a tournament. And we can draw G, D, T as follow.

Definition. A **path** of length k in a graph or digraph is a sequence of distinct vertices $\{v_1, v_2, \ldots, v_{k+1}\}$ such that $(v_i, v_{i+1}) \in E$. A **walk** of length k in a graph or digraph is a sequence of vertices $\{v_1, v_2, \ldots, v_{k+1}\}$ such that $(v_i, v_{i+1}) \in E$. A **loop** is an edge of the form (v, v).

Example. Let G and D be the graph and digraph defined above. Then we have $1,2,3$ is a path and $1,2,1,3$ is a walk and $(2,2)$ is a loop in the graph G. And we have $1,2,3$ is a path and $1,4,3,1$ is a walk and $(2,2)$ is a loop in the digraph D.

Definition. In a graph, **degree** of a vertex v, denoted by $d(v)$, is the number of edge adjacent to v. That is, number of elements in edge set of the form $(\cdot, v)$. For convenience, we say that a loop contribute a vertex degree 1. In a digraph, **out degree** of a vertex v, denoted by $d^+(v)$, is the number of edges of the form $(v, \cdot)$; while **in degree** of a vertex v, denoted by $d^-(v)$, is the number of edges of the form $(\cdot, v)$. For convenience, we say that a loop contribute a vertex out degree 1 and in degree 1.

Example. In the graph G above, we have $d(1) = d(3) = 2$, $d(2) = 4$, and $d(4) = 0$. In the digraph D above, we have $d^+(1) = d^+(2) = 3$, $d^+(3) = 2$, $d^+(4) = 1$ and $d^-(1) = 2$, $d^-(2) = d^-(3) = 3$, $d^-(4) = 1$

Definition. For an $n \times n$ symmetric matrix A, we can associated a graph $\mathcal{G}(A)$ with it. The graph has vertex set $\{1, 2, \ldots, n\}$ and edge set $\{(i,j) : a_{ij} \neq 0\}$. For an $n \times n$ matrix B, we can associated a digraph $\mathcal{G}(A)$ with it. The digraph has vertex set $\{1, 2, \ldots, n\}$ and edge set $\{(i,j) : a_{ij} \neq 0\}$. And the incidence matrix of a graph is the matrix setting $a_{ij} = a_{ji} = 1$ if (i,j) is an edge and $a_{ij} = 0$ for otherwise. And the incidence matrix of a digraph is the matrix setting $a_{ij} = 1$ if (i,j) is an edge and $a_{ij} = 0$ for otherwise. So we have an incidence matrix with dominance relation is actually the incidence matrix of a tournament.

Definition. A **clique**[1] is the maximal set such that each vertex connects to each others. For digraph, v connect to u means $(v,u), (u,v)$ are elements in the edge set.

Note. By induction and some arguement, we can prove that if A is incidence of a graph(digraph) then $(A^k)_{ij}$ is the number of walk from i to j of a graph(digraph).

---

[1]This definition is different from the "clique" in general Graph Theory. In general, a clique means a subset of vertex of some graph such that each vertices are adjacent to each others.