Solutions for exercises/Notes for Linear Algebra Done Right by Sheldon Axler

Toan Quang Pham
mathtangents@gmail.com

Monday 10th September, 2018

Contents

1. Some note before reading the book

1. If the problem does not specify that vector space V is finite-dimensional then it means V can be either finite-dimensional or infinite-dimensional, which is sometimes harder to solve. For example, exercises 24 in 3B.

2. Carefully check whether the theorem/definition holds for real vector space or complex vector space.

2. Terminology

1. There exists a basis of V with respect to which the matrix of T = the matrix of T with respect to basis of V.

2. Eigenvectors $v_1, \ldots, v_m$ corresponding to eigenvalues $\lambda_1, \ldots, \lambda_m$ = Eigenvalues $\lambda_1, \ldots, \lambda_m$ and corresponding eigenvectors $v_1, \ldots, v_m$.

3. Chapter 1 - Vector Spaces

3.1. Excercises 1.B

1. $-(-v) = (-1)(-v) = (-1)((-1)v) = 1v = v$.

2. If $a = 0$ then it's obviously true. If $a \neq 0$ then $v = \left(\frac{1}{a}a\right) v = \frac{1}{a}(av) = \frac{1}{a}0 = 0$.

3. Let $\frac{1}{3}(w - v) = x$ then $3x = w - v$ so $3x + v = (m - v) + v = m$. We can see that x in here is unique, otherwise if there exist such two x, x' then $x = \frac{1}{3}(w - v) = x'$.

4. The empty set is not a vector space because it doesn't contain any additive identity 0 (it doesn't have any vector).

5. Because the collection of objects satisfying the orgininal condition is all vectors in V, same as collection of objects satisfying the new condition.

6. $\mathbf{R} \cup \{\infty\} \cup \{-\infty\}$ is a vector space over $\mathbf{R}$.

3.2. 1.C Subspaces

Problem 3.2.1 (Example 1.35)**.** (a) Obvious.

(b) Let S be set of all continuous real-valued functions on the interval $[0, 1]$. S has additive identity (Example 1.24). If $f, g \in S$ then $f + g, cf$ is continuous real-valued function on the interval $[0, 1]$ for any $c \in \mathbf{F}$.

(c) The idea is similar, if g, f are two differentiable real-valued functions then so are $f + g, cf$.

(d) Let $g, f \in S$ then $g'(2) = f'(2) = b$. According to Notion 1.23 then $f + g \in S$ so $b = (f + g)'(2)$ but $(f + g)(x) = f(x) + g(x)$ so $(f + g)'(x) = f'(x) + g'(x)$ implying $b = (f + g)'(2) = f'(2) + g'(2) = 2b$ or $b = 0$.

(e) Let a_n, b_n be to sequences of S then $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = 0$ so $\lim_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty} ca_n = 0$ which follows $a_n + b_n, ca_n \in S$.

Problem 3.2.2 (Page 19)**.** What are subspaces of $\mathbf{R}^2$?

Answer. $\{(0, 0)\}$ and $\mathbf{R}^2$ are obviously two of them. Are there any other subspaces? Let S be a subspace of $\mathbf{R}^2$ and there is an element $(x_1, y_1) \in S$ with $(x_1, y_1) \neq (0, 0)$. Note that $\lambda(x_1, y_1) \in S$ for each $\lambda \in \mathbf{F}$ so this follows that all points in the line $y = \frac{y_1}{x_1}x$ belong to S. If there doesn't exist a point $(x_2, y_2) \in S$ so that (x_2, y_2) doesn't lie in the line $y = \frac{y_1}{x_1}x$ then we can state that S, set of points lie in $y = \frac{y_1}{x_1}x$, is a subspace of $\mathbf{R}^2$. Hence, since point (x_1, y_1) is an arbitrary point, we deduce that all lines in $\mathbf{R}^2$ through the origin are subspaces of $\mathbf{R}^2$.

Now, what if there exists such $(x_2, y_2) \in S$ then the line $y = \frac{y_2}{x_2}x$ also belongs to S. We will state that if there exists two lines belong to S then $S = \mathbf{R}^2$. Indeed, since $\mathbf{R}^2$ is the union of all lines in $\mathbf{R}^2$ through the origin, it suffices to prove that all lines $y = ax$ for some a belongs to $\mathbf{R}$, given that two lines $y = \frac{y_1}{x_1}x$ and $y = \frac{y_2}{x_2}x$ already belong to S. Let $m = \frac{y_1}{x_1}, n = \frac{y_2}{x_2}$. We know that if $(x_1, y_1), (x_2, y_2) \in S$ then $(x_1 + x_2, y_1 + y_2) \in S$. We pick x_1, x_2 so that $(m - a)x_1 = (a - n)x_2$ then $mx_1 + nx_2 = a(x_1 + x_2)$. We pick $y_1 = mx_1, y_2 = nx_2$ then this follows $(y_3, x_3) = (y_1 + y_2, a(x_1 + y_2)) \in S$ which yields that line $y = ax$ belongs to S. Since this is true for arbitrary a so we deduce that $S = \mathbf{R}^2$.

Thus, all subspaces of $\mathbf{R}^2$ are $\{(0, 0)\}, \mathbf{R}^2$ and all lines in $\mathbf{R}^2$ through the origin. $\square$

Problem 3.2.3 (Page 19)**.** What are subspaces of $\mathbf{R}^3$?

Answer. Similar idea to previous problem. It's obvious that $\{(0, 0, 0)\}, \mathbf{R}^3$ are two subspaces of $\mathbf{R}^3$. Let S be a different subspace.

1. If $(x_1, y_1, z_1) \in S$ then $\lambda(x_1, y_1, z_1) \in S$ implying line ℓ in $\mathbf{R}^3$ through the origin and (x_1, y_1, z_1) belongs to S.

2. If there exists $(x_2, y_2, z_2) \in S$ and (x_2, y_2, z_2) does not lie in the line ℓ on $\mathbf{R}^3$ then the plane P through the origin, (x_1, y_1, z_1) and (x_2, y_2, z_2) belongs to S.

3. If there exists $(x_3, y_3, z_3) \in S$ but not in the plane P then $S = \mathbf{R}^3$.

Thus, $\{0\}, \mathbf{R}^3$, all line in $\mathbf{R}^3$ through the origin and all planes in $\mathbb{R}^3$ through the origin are all subspaces of $\mathbf{R}^3$. $\square$

theo_7:1C:1.45 **Theorem 3.2.4** (Sum of subspaces, page 20) Suppose $U_1, \ldots, U_m$ are subspaces of V. Then $U_1 + U_2 + \ldots + U_m$ is the smallest subspace of V containing $U_1, \ldots, U_m$.

Proof. It's not hard to prove that $U_1 + U_2 + \ldots + U_m$ is subspace of V. Let U be a subspace of V containing $U_1, \ldots, U_m$, we will prove that if $v \in U_1 + \ldots + U_m$ then $v \in U$. Indeed, since $v \in U_1 + \ldots + U_m$ then $v = v_1 + \ldots + v_m$ for $v_i \in U_i$ for all $1 \le i \le m$. On the other hand, $v_i \in U$ since U contains $U_1, \ldots, U_m$ so $v = \sum_{i=1}^m v_i \in U$. Thus, our claim is proven, which yields that U contains $U_1 + \ldots + U_m$.

Next, we will prove that $U_1 + \ldots + U_m$ contains $U_1, \ldots, U_m$. This is not hard since we know that $0 \in U_i$ for all $1 \le i \le m$ so for any $v \in U_i$ then

$$v = \underbrace{0 + \ldots + 0}_{i-1} + v + \underbrace{0 + \ldots + 0}_{m-i} \in U_1 + \ldots + U_m.$$

These two argument follows that $U_1 + \ldots + U_m$ is the smallest subspace of V containing $U_1, \ldots, U_m$. ☐

Theorem 3.2.5 (1.44, Condition of a direct sum, page 23)
Suppose $U_1, U_2, \ldots, U_m$ are subspaces of V. Then $U_1 + \ldots + U_m$ is a direct sum if and only if 0 can only be written as sum $u_1 + \ldots + u_m$ where each $u_i = 0$ in U_i.

Proof. If $U_1 + \ldots + U_m$ is a direct sum then obviously 0 is written uniquely as sum $0 + \ldots + 0$.

If 0 is written uniquely as $0 + 0 + \ldots + 0$. Assume the contrary that sum $U = U_1 + \ldots + U_m$ is not a direct sum, which means there exists $u \in U$ so that $u = v_1 + \ldots + v_m = u_1 + \ldots + u_m$ where $u_i, v_i \in U_i$ and there exists i so $u_i \neq v_i$. This follows

$$\sum_{i=1}^m (v_i - u_i) = \sum_{i=1}^m v_i - \sum_{i=1}^m u_i = u - u = 0.$$

In other words, there is another way to write 0 as $(v_1 - u_1) + \ldots + (v_m - u_m)$ where $v_1 - u_1 \in U_i$ and there exists j so $v_j - u_j \neq 0$, a contradiction. Thus, $U_1 + \ldots + U_m$ is a direct sum. ☐

Theorem 3.2.6 (1.45, Direct sum of two subspaces)
Suppose U, V are two subspaces of V. Then $U + W$ is a direct sum if and only if $U \cap W = \{0\}$.

Proof. If $U + W$ is a direct sum. If there exists $v \in U \cap W$ and $v \neq 0$ then $-v = (-1)v \in U$ so $0 = v - v$, which contradicts to the unique representation of 0 as sum of vectors in U, W.

If $U \cap W = \{0\}$. Assume the contrary that $U + W$ is not a direct sum then $0 = u + v$ with $u \in U, v \in W, u, v \neq 0$. This follows $u = -v \in W$ or $u \in U \cap W$, a contradiction. We are done. ☐

3.3. Exercises 1.C

1. (a), (d) are subspaces of $\mathbf{F}^3$. (b) is not subspace of $\mathbf{F}^3$ because it doesn't contain 0. (c) is not subspace of $\mathbf{F}^3$ because $(0, 1, 1) + (1, 1, 0) = (1, 2, 1) \notin \mathbf{F}^3$.

2. Already did.

3. Let S be set of such functions. Let $f, g \in S$ then $f + g$ is a differentiable real-valued function and $(f + g)'(-1) = f'(-1) + g'(-1) = 3f(2) + 3g(2) = 3(f + g)(2)$. Similarly to λf.

4. Similar to example 1.35.

5. Yes.

6. (a) Since $a^3 = b^3$ if and only if $a = b$ so $\{(a, b, c) \in \mathbf{R}^3 : a^3 = b^3\} = \{(a, b, c) \in \mathbf{R}^3 : a = b\}$. It's not hard to prove that this is subspace of $\mathbf{R}^3$.
(b) No. We see that $(1, 1), \left(1 + i, \frac{-1-\sqrt{3}}{2} + \frac{-1+\sqrt{3}}{2}i\right) \in S$ but $\left(2 + i, \frac{1-\sqrt{3}}{2} + \frac{-1+\sqrt{3}}{2}i\right) \notin S$.

7. Set $\{(a, b) : a, b \in \mathbf{Z}\}$ is not subspace of $\mathbf{R}^2$.

8. Set $U = \{(a, b) \in \mathbf{R}^2 : a = b\} \cup \{(a, b) \in \mathbf{R}^2 : a = 2b\}$ is not subspace of $\mathbf{R}^2$ because $(1, 1), (2, 1) \in U$ but $(3, 2) \notin U$.

9. We consider function $f(x)$ so that $f(x) = x$ for $x \in [0, \sqrt{2})$ and $f(x) = f(x+\sqrt{2})$., function $g(x) = x$ for $x \in [0, \sqrt{3})$ and $g(x) = g(x + \sqrt{3})$ then $f + g$ is not periodic.
Indeed, assume that the function $f(x)+g(x)$ is periodic with length T. Note that if $T/\sqrt{2}$ is not an integer then $f(0) < f(T)$, similar to $T/\sqrt{3}$. Thus, since $T/\sqrt{2}$ and $T/\sqrt{3}$ can't be both integers so $f(0)+g(0) < f(T)+g(T)$, a contradiction. Thus, $f+g$ is non-periodic function. This yields that set of periodic function from $\mathbf{R}$ to $\mathbf{R}$ is not a subspace of $\mathbf{R}^{\mathbf{R}}$.

10. For any $u, v \in U_1 \cap U_2$, then $u + v \in U_1, u + v \in U_2$ so $u + v \in U_1 \cap U_2$. Similar to $\lambda u \in U_1 \cap U_2$.

11. Similar to 10.

12. Assume the contrary, that means there exists two subspaces U_1, U_2 of V so that $U_1 \cup U_2$ is also a subspace and there exists $u_1 \in U_1, u_2 \in U_2, u_1 \notin U_2, u_2 \notin U_1$. We have $u_1, u_2 \in U_1 \cup U_2$ so $u_1 + u_2 \in U_1 \cup U_2$. WLOG, assume $u_1 + u_2 \in U_1$ then $(u_1 + u_2) - u_1 \in U_1$ or $u_2 \in U_1$, a contradiction. Thus, either $U_1 \subseteq U_2$ or $U_2 \subseteq U_1$.

13. Let U_1, U_2, U_3 be three subspaces of V so $U_1 \cup U_2 \cup U_3$ is also a subspace. If there exists two among these three subspaces, WLOG $U_1 \subseteq U_2$, then $U_1 \cup U_2 \cup U_3 = U_2 \cup U_3$ is a subspace if and only if $U_3 \subseteq U_2$ or $U_2 \subseteq U_3$, in other words either U_2 contains U_1, U_3 or U_3 contains U_1, U_2.

    If there doesn't exist any two among three subspaces U_1, U_2, U_3 so that one contains the other, then there exists a vector $v \notin U_2, v \in U_1 \cup U_2 \cup U_3$ which yields $v \in U_1 \cup U_3$. WLOG, $v \in U_1$. Since $U_1 \not\subseteq U_2, U_2 \not\subseteq U_1$ so there exists $u \in U_2, u \notin U_1$. With similar argument to 12, we find $u + v \notin U_2, u + v \notin U_1$ so $u + v \in U_3$. Similar $u + 2v \in U_3$ which we find $(u + 2v) - (u + v) = v \in U_3$. Thus, we find that for any $v \notin U_2$ then $v \in U_1 \cap U_3$. This follows $U_1 \cup U_2 = U_1 \cap U_2$ or $U_1 = U_2$, a contradiction to the assumption.

14. Done

15. $U+U$ is exactly U. From Theorem 1.39, $U+U$ contains U. We also notice that U contains $U+U$ because every $u \in U+U$ then $u \in U$.

16. Yes. As long as U, W are sets of vectors of V.

17. Yes. As long as U_1, U_2, U_3 are sets of vectors of V.

18. Operation of the addition on the subspaces of V has additive identity, which is $U = \{0\}$. Only $\{0\}$ has additive inverse.

19. Let $V = \{(x, y, z) \in \mathbf{R}^3 : x, y, z \in \mathbf{R}\}, U_1 = \{(x, 0, 0) \in \mathbf{R}^3 : x \in \mathbf{R}\}, U_2\{(0, x, 0) \in \mathbf{R}^3 : x \in \mathbf{R}\}$ and $W = \{(x, y, 0) \in \mathbf{R}^3 : x, y \in \mathbf{R}\}$ then $U_1 + W = U_2 + W = W$ but $U_1 \neq U_2$.

20. $W = \{(x, 2x, y, 2y) \in \mathbf{F}^4 : x, y \in \mathbf{F}\}$. It is obvious that $W + U$ is a direct sum. It suffices to prove that $\mathbf{F}^4 = W \oplus U$ or prove that for any $(m, n, p, q) \in \mathbf{F}^4$ there exists $(x, 2x, y, 2y) \in W, (x_1, x_1, y_1, y_1) \in U$ so $(m, n, p, q) = (x, 2x, y, 2y) + (x_1, x_1, y_1, y_1)$. We can choose $x = n - m, x_1 = 2m - n, y = q - p, y_1 = 2p - q$.

21. $W = \{(x, y, z, z, x) \in \mathbf{F}^5 : x, y, z \in \mathbf{F}\}$. Vector 0 can be represented uniquely as $w_1 + u_1$ where $w_1 \in W, u_1 \in U$ and $w_1 = u_1 = 0$. Thus, $W \oplus U$ is a direct sum. It suffices to prove for each $(m, n, p, q, r) \in \mathbf{F}^5$ there is $w = (x, y, z, z, x) \in W, u = (x_1, y_1, x_1 + y_1, x_1 - y_1, 2x_1) \in U$ so $(m, n, p, q, r) = w + u$. We can choose $x_1 = r - m, x = 2m - r, z = m - r + \frac{1}{2}p + \frac{1}{2}q, y_1 = \frac{1}{2}p - \frac{1}{2}q, y = n + \frac{1}{2}q - \frac{1}{2}p$.

22. $W_1 = \{(0, 0, x, 0, 0) \in \mathbf{F}^5 : x \in \mathbf{F}\}, W_2 = \{(0, 0, 0, x, 0) \in \mathbf{F}^5 : x \in \mathbf{F}\}, W_3 = \{(0, 0, 0, 0, x) \in \mathbf{F}^5 : x \in \mathbf{F}\}$.

23. Let's take Exercise 20 as counterexample, $V = \mathbf{F}^4, W = \{(x, x, y, y) \in \mathbf{F}^4 : x, y \in \mathbf{F}\}, U_1 = \{(x, 2x, y, 2y) \in \mathbf{F}^4 : x, y \in \mathbf{F}\}$ and $U_2 = \{(x, 3x, y, 3y) \in \mathbf{F}^4 : x, y \in \mathbf{F}\}$.

24. Since $U_e \cap U_o = \{f(x) = 0\}$ so $U_e + U_o$ is a direct sum. Consider $f(x) \in \mathbf{R}^{\mathbf{R}}$ then $f(x) = f_e(x) + f_o(x)$ where $f_e(x) = \frac{1}{2}[f(x) + f(-x)] \in U_e$ and $f_o(x) = \frac{1}{2}[f(x) - f(-x)] \in U_o$. Thus, $\mathbf{R}^{\mathbf{R}} = U_e \oplus U_o$.

4. Chapter 2 - Finite-dimensional vector spaces

4.1. 2.A Span and linear independence

4.1.1. Main theories

theo_8:2A:2.21 **Theorem 4.1.1** (Linear Dependence Lemma) Suppose $v_1, \dots, v_m$ is linearly dependent list in V. There exists $j \in \{1, 2, \dots, m\}$ such that the following holds:

(a) $v_j \in \operatorname{span}(v_1, \dots, v_{j-1})$.

(b) if the jth term is removed from $v_1, \dots, v_m$, the span of the remaining list equals $\operatorname{span}(v_1, \dots, v_m)$.

theo1 **Theorem 4.1.2** (2.23) In a finite-dimensional vector space, the length of every linearly independent list of vector is less than or equal to the length of every spanning list of vectors.

This theorem is also called as Steinitz exchange lemma or Replacement's theorem.

Proof. Let $v_1, v_2, \dots, v_m$ is linearly independent list and $u_1, u_2, \dots, u_n$ is spanning list of vectors. It suffices to prove for the case that $u_1, u_2, \dots, u_n$ is also linearly independent, otherwise if $u_1, u_2, \dots, u_n$ is linearly dependent, according to Linear Dependence Lemma (2.21), we can take out some vectors from the list without changing the span and bring back to linearly independent list. Assume the contrary that $n < m$ and from this assumption, we will prove that list $v_1, \dots, v_m$ is linearly dependent.

Since $u_1, u_2, \dots, u_n$ spans V and is also linearly independent so for each v_i, there exists unique $c_{i1}, c_{i2}, \dots, c_{in} \in \mathbf{F}$ so $v_i = c_{i1}u_1 + c_{i2}u_2 + \dots + c_{in}u_n$. Hence, for $a_1, \dots, a_m \in \mathbf{F}$, we have

$$a_1v_1 + a_2v_2 + \dots + a_mv_m = u_1 \sum_{i=1}^{m} a_ic_{i1} + u_2 \sum_{i=1}^{m} a_ic_{i2} + \dots + u_n \sum_{i=1}^{m} a_ic_{in}.$$

We will show that there exists $a_1, \dots, a_m$ not all 0 so $a_1v_1 + \dots + a_mv_m = 0$. Since $u_1, \dots, u_n$ is linearly independent so form the above, we follow

$$a_1v_1 + \dots + a_mv_m = 0 \iff \begin{cases} a_1c_{11} + a_2c_{21} + \dots + a_mc_{m1} = 0, \\ a_2c_{12} + a_2c_{22} + \dots + a_mc_{m2} = 0, \\ \qquad \dots \\ a_mc_{1n} + a_2c_{2n} + \dots + a_mc_{mn} = 0. \end{cases} \tag{1}$$ eq1

This system of equation has m variable and n equation with $n < m$. Here is an algorithm in finding non-zero solution for $a_1, \dots, a_m$. Number the equations in (1) from top to bottom as $(11), (12), \dots, (1n)$.

1. For $1 \le i \le m - 1$, get rid of a_m in $(1i)$ by subtracting to $(1(i+1))$ times some constant. By doing this, we obtain a system of equations where the first $n-1$ equations only contain $m-1$ variables, the last equation contains m variables.

2. For $1 \le i \le m-2$, get rid of a_m in $(1i)$ by subtracting to $(1(i+1))$ times some constant. We obtain an equivalent system of equations where the first $n-2$ equations contain $m-2$ variables, equation $(1(n-1))$ contains $m-1$ variables and and equation $(1n)$ contains m variables.

3. Keep doing this until we get an system of equation so that the $(1i)$ equation contains $m-n+i$ variables.

4. Starting from equation (11) with $m-n+1$ variables $a_1, \dots, a_{m-n+1}$, we can pick $a_1, \dots, a_{m+n-1}$ so that it satisfies equation (11) and $a_1, \dots, a_{m-n+1}$ not all 0, which is possible since $m \ge n+1$.

5. Come to (12) and pick a_{m-n+2}, to (13) and pick a_{m-n+3} until to $(1n)$ and pick a_m.

Thus, there exists $a_1, \dots a_m$ not all 0 so that $a_1v_1 + \dots + a_mv_m = 0$, a contradiction since $v_1, \dots, v_m$ is linearly independent. Hence, we must have $n \ge m$. We are done. □

See Exercises 12, 13, 14, 17 for applications of this theorem.

Theorem 4.1.3 (2.26 Finite dimensional subspaces) Every subspace of a finite-dimensional vector space is finite-dimensional.

Proof. Let $V = \text{span}(v_1, v_2, \dots, v_m)$ be a finite-dimensional vector space and U be its subspace. Among vectors in the spanning list $v_1, v_2, \dots, v_m$, let $u_1, u_2, \dots u_n$ be vectors that are in subspace U, $k_1, k_2, \dots, k_{m-n}$ be vectors that are not in U. If $U = \text{span}(u_1, u_2, \dots, u_n)$ then we are done. If not, then there exists $\ell \in U$ so $\ell \notin \text{span}(u_1, u_2, \dots, u_n)$. Hence, since $\ell \in \text{span}(v_1, \dots, v_m)$ so

$$\ell = \sum_{i=1}^{n} u_i a_i + \sum_{j=1}^{m-n} k_i b_i$$

for $a_i, b_i \in \mathbf{F}$ and there exists b_i $(1 \le i \le m-n)$ so that $b_i \ne 0$. Since $u_i, \ell \in U$ so that means $h_1 = \sum_{i=1}^{m-n} k_i b_i \in U$ and also in $\text{span}(k_1, \dots, k_{m-n})$ and $h_1 \ne 0$. We add h_1 to the list $h_1, u_1, u_2, \dots, u_n$ in U. Similarly, we compare U and $\text{span}(h_1, u_1, \dots, u_n)$ and may add another $h_2 \in \text{span}(k_1, k_2, \dots, k_{m-n})$ onto the list. This process will keep going until we obtain a spanning list of U. We will prove that after adding at most $m-n$ such h_i, we can get $U = \text{span}(h_1, h_2, \dots, h_j, u_1, \dots, u_n)$ with $j \le m-n$.

Indeed, assume the contrary that after adding at least $m-n$ such $h_i \in U$ so $h_i \ne 0$, we still can't obtain a list of vectors that spans U. That means there exists $\ell \in U$ and $\ell \notin \text{span}(h_1, \dots, h_{m-n}, u_1, \dots, u_n)$. Similarly, we can construct another , $h_{m+n-1} \ne 0, h_{m-n+1} \in U$ and $h_{m-n+1} \in \text{span}(k_1, \dots, k_{m-n})$ based on ℓ. Let $h_i = c_{i1}k_1 + c_{i2}k_2 + \dots + c_{i,m-n}k_{m-n}$ then we have

$$\begin{aligned}
&a_1h_1 + \ldots + a_{m-n+1}h_{m-n+1} = k_1 \\
\iff &\begin{cases} a_1c_{1,1} + a_2c_{2,1} + \ldots + a_{m-n+1}c_{m-n+1,1} = 1, \\ a_1c_{1,2} + a_2c_{2,2} + \ldots + a_{m-n+1}c_{m-n+1,2} = 0, \\ \qquad \ldots \\ a_1c_{1,m-n} + \ldots + a_{m-n+1}c_{m-n+1,m-n} = 0. \end{cases}
\end{aligned}$$

This is a system of $m-n$ equations with $m-n+1$ variables $a_1, \ldots, a_{m-n+1} \in \mathbf{F}$ so similarly to Theorem 4.1.2's proof, there are infinite number of a_i so $a_1h_1 + \ldots + a_{m-n+1}h_{m-n+1} = k_1$. Since $h_i \in U$ so $k_1 \in U$, a contradiction. Thus, there are at most $m-n$ such $h_i \in U$ $(1 \leq i \leq j)$ so that there does not exist $\ell \in U$ but $\ell \notin \text{span}(h_1, \ldots, h_j, u_1, \ldots, u_n)$. Note that $\text{span}(h_1, \ldots, h_j, u_1, \ldots, u_n) \subseteq U$ so this can only mean $U = \text{span}(h_1, \ldots, h_j, u_1, \ldots, u_n)$. Thus, U is a finite-dimensional vector space. $\square$

Remark 4.1.4. Both my proofs for the two theorems use the result that a system of m equations with n variables in $\mathbf{R}$ with $m < n$ has infinite solution in $\mathbf{R}$.

4.1.2. Important/Interesting results from Exercise 2.A

Corollary 4.1.5 (Example 2.20) If some vector in a list of vectors in V is a linear combination of the other vectors, then the list is linearly dependent.

Proof. List $v_1, \ldots, v_m$ has $v_1 = \sum_i a_iv_i$ then $v_1 - \sum_i a_iv_i = 0$, or there exists $x_1, \ldots, x_m$ not all 0 so $\sum_{i=1}^m x_iv_i = 0$. Hence, $v_1, \ldots, v_m$ is linearly dependent. $\square$

This is another way to check if some list is linearly independent or not. See Exercises 10 (2A) for application. The reverse is also true.

Corollary 4.1.6

(Exercise 11, 2A) Suppose $v_1, \ldots, v_m$ is linearly independent in V and $w \in V$ then $v_1, v_2, \ldots, v_m, w$ is linearly independent if and only if $w \notin \text{span}(v_1, v_2, \ldots, v_m)$.

This is a corollary implying from Linear Dependence Lemma (2.21). With this, we can construct a spanning list for any infinite-dimensional vector space. Also with this, we can prove Exercise 14, which proves existence of a sequence of vectors in infinite-dimensional vector space so that for any positive integer m, the first m vectors in the sequence is linearly independent. See Exercise 15, 16 for applications for Exercise 14.

4.2. Exercises 2.A

1. For any $v \in V$ then $v = a_1v_1 + a_2v_2 + a_3v_3 + a_4v_4$ for $a_i \in \mathbf{F}$. Hence,

$$\begin{aligned}
v &= a_1(v_1 - v_2) + (a_1 + a_2)(v_2 - v_3) + (a_1 + a_2 + a_3)(v_3 - v_4) + (a_1 + a_2 + a_3 + a_4)v_4, \\
&= b_1(v_1 - v_2) + b_2(v_2 - v_3) + b_3(v_3 - v_4) + b_4v_4.
\end{aligned}$$

Thus, $v \in \text{span}(v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4)$. We follow $V \subseteq \text{span}(v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4)$. On the other hand, since $v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$ so for any $u \in \text{span}(v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4)$ then $u \in V$. We find $\text{span}(v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4 \subseteq V$. Thus, $v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$ spans V.

2. Example 2.18: (a) True for $v \neq 0$. If $v = 0$ then for any $a \in \mathbf{F}$ we always have $av = 0$, so list of one vector 0 is not linearly independent.
(b) Consider $u, v \in V$. We have $au + bv = 0 \iff au = -bv$ so in order to have only one choice of (a, b) to be $(0, 0)$ then u can't be scalar multiple of v.
(c),(d) True.

3. It suffices to find t so there are $x, y, z \in \mathbf{R}$ other than $x = y = z = 0$ so $x \cdot (3, 1, 4) + y \cdot (2, -3, 5) + z \cdot (5, 9, t) = 0$ or $\begin{cases} 3x + 2y + 5z = 0, \\ x - 3y + 9z = 0, \\ 4x + 5y + tz = 0. \end{cases}$ Pick $t = 2$ then we find $(x, y, z) = (-3a, 2a, a)$ for any $a \in \mathbf{R}$. This follows $(3, 1, 4), (2, -3, 5), (5, 9, t)$ is not linearly independent in $\mathbf{R}^3$.

4. Similarly to 3, let say if $x \cdot (2, 3, 1) + y \cdot (1, -1, 2) + z \cdot (7, 3, c) = 0$ for $x, y, z \in \mathbf{R}$ then $\begin{cases} 2x + y + 7z = 0, \\ 3x - y + 3z = 0, \\ x + 2y + cz = 0 \end{cases}$ From two above equations, we find $x = -2z, y = -3$ and substitute into the third equation to get $(c - 8)z = 0$. This has at least two solutions of z if and only if $c = 8$. Done.

5. (a) The only solution $x, y \in \mathbf{R}$ so $x(1 + i) + y(1 - i) = 0$ is $x = y = 0$. Hence list $1 + i, 1 - i$ is linearly independent in vertor space $\mathbf{C}$ over $\mathbf{R}$.
(b) Beside $x = y = 0$ then $x = i, y = 1$ also satisfies $x(1 + i) + y(1 - i) = 0$. Hence, $1 + i, 1 - i$ is linearly dependent in vector space $\mathbf{C}$ over $\mathbf{C}$.

exer:2A:6 6. Consider $a, b, c, d \in \mathbf{F}$ so $a(v_1 - v_2) + b(v_2 - v_3) + c(v_3 - v_4) + dv_4 = 0$, which is equivalent to $av_1 + (b - a)v_2 + (c - b)v_3 + (d - c)v_4 = 0$. Since v_1, v_2, v_3, v_4 is linearly dependent in V so this follows $a = b - a = c - b = d - c = 0$ or $a = b = c = d = 0$ and a, b, c, d are uniquely determined. Hence, $v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$ is linearly independent in V.

exer:2A:7 7. True. Let $a_1(5v_1 - 4v_2) + a_2v_2 + a_3v_3 + \ldots + a_mv_m = 0$ or $5a_1v_1 + (a_2 - 4a_1)v_2 + a_3v_3 + \ldots + a_mv_m = 0$, which implies $5a_1 = a_2 - 4a_1 = a_3 = \ldots = a_m = 0$ or $a_1 = a_2 = \ldots = a_m$.

8. Consider $a_1\lambda v_1 + a_2\lambda v_2 + \ldots + a_m\lambda v_m = 0$, which implies $a_1\lambda = a_2\lambda = \ldots = a_m\lambda = 0$ since $a_1, a_2, \ldots, a_m$ is linear independent. Thus, $a_1 = a_2 = \ldots = a_m$ so $\lambda v_1, \lambda v_2, \ldots, \lambda v_m$ is linearly independent.

9. Not true. Let $v_1, v_2, \ldots, v_m$ be $(1, 0, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, \ldots, 0, 1)$, respectively and let $w_3, \ldots, w_m$ be $(0, 0, 2, 0, \ldots, 0), \ldots, (0, \ldots, 0, 2)$ and $w_1 = (0, -1, 0, \ldots, 0)$ and

$w_2 = (-1, 0, \ldots, 0)$. Therefore, for $i \geq 3$ then $w_i + v_i = (\underbrace{0, \ldots, 0}_{i-1}, i, 0, \ldots, 0)$ and $w_1 + v_1 = (1, -1, 0, \ldots, 0), w_2 + v_2 = (-1, 1, 0, \ldots, 0)$. It's not hard to see $(w_1 + v_1) + (w_2 + v_2) = 0$ so $w_1 + v_1, w_2 + v_2, \ldots, w_m + v_m$ is linearly dependent.

10. If $v_1 + w, v_2 + w, \ldots, v_m + w$ is linearly dependent then according to Linear Dependence Lemma (2.21) there exists $j \in \{1, 2, \ldots, m\}$ so that $v_j + m \in \operatorname{span}(v_1 + m, \ldots, v_{j-1} + m)$. Hence, $v_j + m = a_1(v_1 + m) + \ldots + a_{j-1}(v_{j-1} + m)$ or

$$(a_1 + \ldots + a_{j-1} - 1)m = v_j - a_1 v_1 - a_2 v_2 - \ldots - a_{j-1} v_{j-1}$$

If $a_1 + \ldots + a_{j-1} - 1 = 0$ then $v_j - a_1 v_1 - a_2 v_2 - \ldots - a_{j-1} v_{j-1} = 0$. If $j = 1$ then that means $v_j = 0$, so $v_1, v_2, \ldots, v_m$ is linearly dependent, a contradiction. If $j \geq 2$ then $v_j = a_1 v_1 + \ldots + a_{j-1} v_{j-1}$ which also follows that $v_1, \ldots, v_m$ is linearly dependent, a contradiction.

Thus, we must have $a_1 + \ldots + a_{j-1} - 1 \neq 0$. Therefore, since $RHS \in \operatorname{span}(v_1, \ldots, v_m)$ so $m \in \operatorname{span}(v_1, \ldots, v_m)$.

exer:2A:11 11. If $v_1, \ldots, v_m, w$ is linearly independent then $w \notin \operatorname{span}(v_1, \ldots, v_m)$ according to Corollary 4.1.5. The reverse, if $w \notin \operatorname{span}(v_1, \ldots, v_m)$ then assume the contrary, if $v_1, \ldots, v_m, w$ is linearly dependent then it will lead to a contradiction according to Linear Dependence Lemma (2.21). Done.

exer:2A:12 12. Because $\mathcal{P}_4(\mathbf{F}) = \operatorname{span}(1, x, x^2, x^3, x^4)$ and from Theorem 4.1.2 (2.23) length of any linearly independent list is less than length of a spanning list $1, x, x^2, x^3, x^4$, which is 5. That explains why we can't have 6 polynomials that is linearly independent in $\mathcal{P}_4(\mathbf{F})$.

exer:2A:13 13. Because $1, x, x^2, x^3, x^4$ is a linearly independent list of length 5 in $\mathcal{P}_4(\mathbf{F})$ and according to Theorem 4.1.2 (2.23) length of any spanning list is greater than or equal to length of a linearly independent list, which is 5. Thus, we can't have four polynomials that spans $\mathcal{P}_4(\mathbf{F})$.

exer:2A:14 14. Prove that V is infinite-dimensional if and only if there is a sequence $v_1, v_2, \ldots$ of vectors in V such that $v_1, \ldots, v_m$ is linearly independent for every positive integer m.

Proof. If V is infinite-dimensional: Pick $v_1 \in V$, pick v_i so $v_i \in V, v_i \notin \operatorname{span}(v_1, \ldots, v_{i-1})$. Such v_i exists, otherwise any $v \in V$ must also in $\operatorname{span}(v_1, \ldots, v_{i-1})$ so $V \subseteq \operatorname{span}(v_1, \ldots, v_{i-1})$ and we also have $\operatorname{span}(v_1, \ldots, v_{i-1}) \subseteq V$ so a finite list of vectors span V, a contradiction. Note that since $v_1, \ldots, v_{i-1}$ is linearly independent and $v_i \notin \operatorname{span}(v_1, \ldots, v_{i-1})$ so by Corollary 4.1.6 then $v_1, \ldots, v_i$ is linearly independent. By continuing the process, we can construct such sequence of vectors. □

If there exists such sequence of vectors in V: Assume the contrary that V is finite-dimensional then there exists finite spanning list in V with length ℓ. However, from existence of such sequence, we can find in V a linearly independent list with arbitrary large length, which contradicts to Theorem 4.1.2 that length of any linearly independent list less than or equal to length of spanning list in a finite-dimensional vector space. We are done.

exer:2A:15 15. With $\mathbf{F}^\infty$, consider the sequence $(1,0,\dots),(0,1,0,\dots),(0,0,1,0,\dots),\dots$ then from Exercise 14, we follow $\mathbf{F}^\infty$ is infinite-dimensional.

exer:2A:16 16. Consider sequence of continuous real-valued functions defined on $[0,1]$: $1,x,x^2,x^3,\dots$ and we are done.

exer:2A:17 17. Nice problem. Since $p_j(2)=0$ for each j so that means $x-2$ divides each p_j. Let $q_i=\frac{p_i}{x-2}$ then $q_0,q_1,\dots,q_m$ are polynomials in $\mathcal{P}_{m-1}(\mathbf{F})$. If $q_0,q_1,\dots,q_m$ is linear independent but it has length $m+1$ larger than length m of spanning list in $\mathcal{P}_{m-1}(\mathbf{F})$, a contradiction according to Theorem 4.1.2. Thus, $q_0,q_1,\dots,q_m$ is linear dependent. That means there exists $a_i\in\mathbf{F}$ not all 0 so $\sum_{i=0}^m a_iq_i=0$. This follows $(x-2)\sum_{i=0}^m a_iq_i=0$ or $\sum_{i=0}^m a_ip_i=0$. Hence, $p_0,\dots,p_m$ is linearly dependent.

4.3. 2.B Bases

theo_3:2B:2.31 **Theorem 4.3.1** (2.31) Every spanning list in a vector space can be reduced to a basis of the vector space.

theo_4:2B:2.32 **Theorem 4.3.2** (2.32) Every finite-dimensional vector has a basis.

theo_5:2B:2.33 **Theorem 4.3.3** (2.33) Every linearly independent list of vectors in a finite-dimensional vector space can be extended to a basis of the vector space.

theo_6:2B:2.34 **Theorem 4.3.4** (2.34) Suppose V is finite-dimensional and U is subspace of V. Then there is a subspace W of V such that $V=U\oplus W$.

Proof. According to Theorem 4.1.3 then U is also finite-dimensional. Hence, according to Theorem 4.3.2 then U has a basis $v_1,\dots,v_m$. Since $v_1,\dots,v_m$ is linearly independent so according to Theorem 4.3.3, there exists vectors $w_1,w_2,\dots,w_n$ so $v_1,\dots,v_m,w_1,\dots,w_n$ is a basis of V Let $W=\operatorname{span}(w_1,\dots,w_n)$ then $V=U\oplus W$. □

Exercise 8 chapter 2B gives another rephrasing of theorem 4.3.4, saying that if there exists subspaces U,W of V so $V=U\oplus W$ and each has basis $\{u_1,\dots,u_m\}$ and $\{w_1,\dots,w_n\}$ respectively then $u_1,\dots,u_m,w_1,\dots,w_n$ is a basis of V.

4.4. Addition, scalar multiplication of specific list of vectors

The book doesn't seem to mention this. I just realised this while i was doing the exercises. It makes the writing much easier if I write a generalisation in here so I can refer to it later in the exercises. It basically says spanning list/basis/ linearly independent list has addition and scalar multiplication properties.

prop:1 **Proposition 4.4.1** Given $(v_1, v_2, \ldots, v_m)$ that is a basis/spanning list/linearly independent list then $(v_1, v_2, \ldots, v_i + w, v_{i+1}, \ldots, v_m)$ is also a basis/spanning list/linearly independent list with $w = \sum_{j=1}^m b_j v_j$ with $b_i \neq -1$ and $b_j \in \mathbf{F}$ for all $1 \le j \le m$..

Proof. It's not hard to see that if $(v_1, \ldots, v_m)$ is a spanning list then so is $(v_1, \ldots, v_i + w, v_{i+1}, \ldots, v_m)$.

If $(v_1, \ldots, v_m)$ is linearly independent, we consider $a_i \in \mathbf{F}$ so $a_1v_1 + a_2v_2 + \ldots + a_i(v_i + w) + \ldots + a_mv_m = 0$ or

$$(a_1 + a_ib_1)v_1 + (a_2 + a_ib_2)v_2 + \ldots + (a_i + a_ib_i)v_i + \ldots + (a_m + a_ib_m)v_m = 0.$$

This follows $a_j + a_ib_j = 0$ for all $1 \le j \le m$. Consider $a_i + a_ib_i = 0$ or $a_i(1 + b_i) = 0$, since $b_i \neq -1$ so $a_i = 0$. We also have $a_j + a_ib_j = 0$ for all $1j \le m$ so we find $a_j = 0$ for all $1 \le j \le m$. Hence, $(v_1, v_2, \ldots, v_i + w, \ldots, v_m)$ is linearly independent. □

Using this proposition, we can solve exercises 6, 7 in chapter 2A. We can construct a new basis from known basis of a vector, such as exercises 3, 4, 5, 6 in chapter 2B.

4.5. Exercises 2B

1. If $(v_1, \ldots, v_m)$ is a basis in V then so is $(v_1, \ldots, v_{m-1}, v_m + v_1)$ according to proposition 4.4.1. Hence in order for V to have only one basis, we must have $v_m = v_m + v_1$ or $v_1 = 0$. Since vector 0 in the list so in order for this list to be linearly independent, the list must contain 0 only. Thus, this follows $\{0\}$ is the only vector space that has only one basis.

2. Not hard to do.

exer:2B:3 3. (a) $U = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbf{R}^5 : x_1 = 3x_2,\ x_3 = 7x_4\}$. Basis of U is $\{(3, 1, 0, 0, 0), (0, 0, 7, 1, 0), (0, 0, 0,$

   (b) Add two more to basis of U which are $(0, 1, 0, 0, 0)$ and $(0, 0, 0, 1, 0)$ we will get a basis of $\mathbf{R}^5$.

   (c) According to proof of theorem 4.3.4, pick $W = \operatorname{span}((0, 1, 0, 0, 0), (0, 0, 0, 1, 0))$. then $\mathbf{R}^5 = U \oplus W$.

exer:2B:4 4. (a) Basis of $U = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbf{C}^5 : 6z_1 = z_2,\ z_3 + 2z_4 + 3z_5 = 0\}$ can be $(1, 6, 0, 0, 0), (0, 0, 4, 1, -2), (0, 0, 7, 1, -3)$. We need to prove that this list is a basis, i.e. there exists uniquely $m, n \in \mathbf{C}$ so $m(4, 1, -2) + n(7, 1, -3) = (z_3, z_4, z_5)$ for any $z_3, z_4, z_5 \in \mathbf{C}$ so $z_3 + 2z_4 - 3z_5 = 0$. Note that $4 + 2 \cdot 1 + 3 \cdot (-2) = 0$ and $7 + 2 \cdot 1 + 3 \cdot (-3) = 0$ so as long as we determine m, n for z_3, z_4 then that m, n also works for z_5. Hence, it deduces to $m(4, 1) + n(7, 1) = (z_3, z_4)$ or $4m + 7n = z_3$ and $m + n = z_4$. We have system of equations $\begin{cases} 4\operatorname{Re}(m) + 7\operatorname{Re}(n) = \operatorname{Re}(z_3) \\ \operatorname{Re}(m) + \operatorname{Re}(n) = \operatorname{Re}(z_4) \end{cases}$ and

$\begin{cases} 4\,\mathrm{Im}(m) + 7\,\mathrm{Im}(n) = \mathrm{Im}(z_3) \\ \mathrm{Im}(m) + \mathrm{Im}(n) = \mathrm{Im}(z_4) \end{cases}$. Hence, m, n can be uniquely determined. Thus, the list is a basis.

(b) Proposition 4.4.1 can construct a basis of $\mathbf{C}^5$ than contains three vectors in (a). Let $v_i = \left(\underbrace{0, 0, \dots, 0}_{i-1}, 1, 0, \dots, 0\right)$ then v_1, v_2, v_3, v_4, v_5 is a basis of $\mathbf{C}^5$ then according proposition 4.4.1 so does $\{v_1 + 6v_2, v_2, 4v_3 + v_4 - 2v_5, v_4, v_5\} = \{u_1, v_2, u_3, v_4, v_5\}$, which is

$$(1, 6, 0, 0, 0), (0, 1, 0, 0, 0), (0, 0, 4, 1, -2), (0, 0, 0, 1, 0), (0, 0, 0, 0, 1).$$

From here we find $\left\{u_1, v_2, u_3, \frac{7}{4}u_3 - \frac{3}{4}v_4 + \frac{1}{2}v_5, v_5\right\}$ is also a basis, or

$$(1, 6, 0, 0, 0), (0, 1, 0, 0, 0), (0, 0, 4, 1, -2), (0, 0, 7, 1, -3), (0, 0, 0, 0, 1).$$

(c) According to proof of theorem 4.3.4, pick $W = \mathrm{span}((0, 1, 0, 0, 0), (0, 0, 0, 0, 1))$ and we are done.

exer:2B:5 5. True. Note that $1, x, x^2, x^3$ is a basis of $\mathcal{P}_3(\mathbf{F})$ so according proposition 4.4.1 we find $1, x, x^2 + x^3, x^3$ is also a basis of $\mathcal{P}_3(\mathbf{F})$ and one of the vectors has degree 2.

exer:2B:6 6. True according to proposition 4.4.1.

7. It's not true. A counterexample: Pick $V = \mathbf{F}^4$ and $U = \{(x_1, x_2, x_3, x_4) \in \mathbf{F}^4 : x_2 = 2x_1\}$. Note that $v_1 = (0, 0, 0, 1), v_2 = (0, 0, 1, 0) \in U$ and $v_4 = (1, 0, 0, 0), v_3 = (0, 1, 0, 0) \notin U$ and $(0, 0, 0, 1), (0, 0, 1, 0)$ is not a basis of U (it does not span U).

exer:2B:8 8. Since $V = U \oplus W$ so every $v \in V$ can be represented uniquely as $v = ru + sw$ with $u \in U, w \in W$. Since $u_1, \dots, u_m$ is a basis of U so u can be represented uniquely as $\sum_{i=1}^m a_i u_i$. Similarly, w can be represented uniquely as $w = \sum_{i=1}^n b_i w_i$. Hence, v can be represented uniquely as a linear combination of $u_1, \dots, u_m, w_1, \dots, w_n$. Hence, $u_1, \dots, u_m, w_1, \dots, w_n$ is a basis of V.

4.6. 2.C Dimension

theo_9:2C:2.35 **Theorem 4.6.1** (2.35) Any two bases of a finite-dimensional vector space the a same length.

Proof. This is true according to theorem 4.1.2. By considering a basis $v_1, \dots, v_m$ in finite-dimensional vector space V then $v_1, \dots, v_m$ is a list length m which is both linearly independent and spanning V. Other bases are also linearly independent lists so they must have length less than length of any spanning list, which is m. They are also spanning lists so their lengh must be greater than length of any independent list, which is m. Hence, other bases must also have length m. □

Definition 4.6.2. The dimension of a finite-dimensional vector space is the length of any basis of the vector space. The dimension of V (if V is finite-dimensional) is denoted by $\dim V$.

theo_10:2C:2.38 **Theorem 4.6.3** (2.38) If V is finite-dimensional and U is a subspace of V then $\dim U \le \dim V$.

Proof. This is true according to theorem 4.3.3, since basis $u_1, \dots, u_m$ of U is linearly independent so there exists $w_1, \dots, w_n \in V$ so $u_1, \dots, u_m, w_1, \dots, w_n$ is a basis of V. Therefore, $\dim U \le \dim V$. □

theo_11:2C:2.39 **Theorem 4.6.4** (2.39) Suppose V is finite-dimensional. Then every linearly independent list of vectors in V with length $\dim V$ is a basis of V.

Proof. According to theorem 4.3.3, every linearly independent list can be extended to a basis. Since the list has length $\dim V$, that means no vector need to be added to the list to create a basis, i.e. the list is a basis. □

theo_12:2C:2.42 **Theorem 4.6.5** (2.42) Suppose V is finite-dimensional. Then every spanning list of vectors in V with length $\dim V$ is a basis of V.

theo_13:2C:2.43 **Theorem 4.6.6** (Dimension of a sum, 2.43) If U_1 and U_2 are subspaces of a finite-dimensional vector space, then

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim (U_1 \cap U_2).$$

Proof. Note that $U_1 \cap U_2$ is a subspace of U_1, U_2. Hence, if we let $v_1, \dots, v_m$ be basis of vector space $U_1 \cap U_2$ then according to theorem 4.3.3, there exists $w_1, \dots, w_n \in U_1$ so $v_1, \dots, v_m, w_1, \dots, w_n$ is basis of U_1 and there exists $u_1, \dots, u_l \in U_2$ so $v_1, \dots, v_m, u_1, \dots, u_l$ is basis of U_2. Thus, this follows the right hand side of the identity will be $m+n+l$. Hence, it suffices to prove $\dim(U_1+U_2)=m+n+l$. Indeed, we will prove that

$$v_1,\dots,v_m,u_1,\dots,u_l,w_1,\dots,w_n \tag{2}$$

eq2:2C:2.43

is a basis of U_1+U_2. This list obviously spans U_1+U_2. Before, we go and prove the above is is linearly independent, notice that form exercise 8 2B then $U_1=(U\cap U_2)\oplus S$ where $S=\text{span}(w_1,\dots,w_n)$ so this follows $S\cap(U_1\cap U_2)=\{0\}$ from theorem Direct sum of two subspaces 3.2.4.

Now, coming back to (2), assume the list is linearly depdendent, and note that $v_1,\dots,v_m,u_1,\dots,u_l$ is linearly independent so from Linear Dependence Lemma 4.1.1 we find that there exists $1\le j\le n$ so $w_i\in\text{span}(v_1,\dots,v_m,u_1,\dots,u_l,w_1,\dots,w_{j-1})$, or there exists $v\in S, v\neq 0$ so $v\in U_2$. Note that $S\subset U_1$ so from this we find $v\in U_1\cap U_2$. Hence, $v\in(S\cap(U_1\cap U_2))$ and $v\neq 0$, a contradiction with above argument. Thus, the list is linearly independent. We obtain list (2) is a basis of U_1+U_2 so $\dim\,(U_1+U_2)=m+n+l=\dim\,U_1+\dim\,U_1-\dim\,(U_1+U_2)$. $\square$

4.7. Exercises 2C

1. Let $v_1,\dots,v_m$ be basis of U then this list is linearly independent which has length equal to $\dim V$ so from theorem 4.6.4 the list is a basis of V, therefore, $U=V$.

2. We have $\dim\mathbf{R}^2=2$ and from theorem 4.6.3 then if U is subspace of $\mathbf{R}^2$ then $\dim U\le 2$. If $\dim U=0$ then $U=\{0\}$, if $\dim U=1$ then U is $U=\text{span}((x,y))$ which is a line in $\mathbf{R}^2$ through the origin and $(x,y)\neq 0$. If $\dim U=2$ then $U=\mathbf{R}^2$.

3. Similarly, we find $\dim U\le 3$ for U a subspace of $\mathbf{R}^3$. If $\dim U=0$ then $U=\{0\}$, if $\dim U=1$ then U is set of points in line in $\mathbb{R}^3$ through the origin. If $\dim U=2$ then $U=\text{span}((x_1,y_1,z_1),(x_2,y_1,z_2))$ so $(x_1,y_1,z_1)\neq k(x_2,y_2,x_2)$ with $k\mathbf{R}$, i.e. they are not in the same line through the origin. This follows U is the plane through the origin, (x_1,y_1,z_1) and (x_2,y_2,z_2).

4. a) With $U=\{p\in\mathcal{P}_4(\mathbf{F}) : p(6)=0\}$ then $x-6,(x-6)x,(x-6)x^2,(x-6)x^3$ is a basis of subspace U of $\mathcal{P}_4(\mathbf{F})$. Indeed, it's obvious that the list spans U. The list is also linearly independent since it doesn't exist a,b,c,d so $a(x-6)+b(x-6)x+c(x-6)x^2+d(x-6)x^3=0$ for all $x\in\mathbf{R}$.

   b) $1,x-6,(x-6)x,(x-6)x^2,(x-6)x^3$ is a basis of $\mathcal{P}_4(\mathbf{F})$ since this is spans $\mathcal{P}_4(\mathbf{F})$ and it has length $5=\dim\mathcal{P}_4(\mathbf{F})$.

   c) According to theorem 4.3.4 then $W=\text{span}\,(1)$.

5. a) If $U=\{p\in\mathcal{P}_4(\mathbf{F}) : p''(6)=0\}$ then $1,x,(x-6)^3,(x-6)^4$ is a basis of U. This is is obviously linearly independent. Since $\dim\mathcal{P}_4(\mathbf{F})=5$ so from theorem 4.6.3 then $\dim U\le 5$ and if $\dim U=5$ then $U=\mathcal{P}_4(\mathbf{F})$ according to theorem 4.6.4, which is a contradiction since $x^2\notin U, x^2\in\mathcal{P}_4(\mathbf{F})$. Thus, $\dim U=4$ so again from 4.6.4 we have $1,x,(x-6)^3,(x-6)^4$ is a basis of U.

b) Add x^2 to $1, x, (x-6)^3, (x-6)^4$ we get a basis of $\mathcal{P}_4(\mathbf{F})$ because this list is linearly independent and as length of 5.

c) According to theorem 4.3.4, we find $W = \text{span }(x^2)$.

exer:2C:6

6. a) With $U = \{p \in \mathcal{P}_4(\mathbf{F}) \ : \ p(2) = p(5)\}$ then $1, (x-2)(x-5), (x-2)(x-5)x, (x-2)(x-5)x^2$ is a basis of U. Indeed, this is linearly independent and similarly to above argument, we find dim U ≤ 4 so we are done.

   b) Add x to the list and we obtain a linearly independent list with length of 5 so it is a basis of $\mathcal{P}_4(\mathbf{F})$.

   c) According to theorem 4.3.4, we find $W = \text{span }(x)$.

7. a) With $U = \{p \in \mathcal{P}_4(\mathbf{F}) \ : \ p(2) = p(5) = p(6)\}$ then $1, (x-2)(x-5)(x-6), (x-2)(x-5)(x-6)x$ is a basis of U. Indeed, this is is linearly independent and it is a subspace of $U_2 = \{p \in \mathcal{P}_4(\mathcal{F}) \ : \ p(2) = p(5)\}$. From exercise 6 2C we know $\dim U_2 = 4$ so $\dim U \leq 4$. If $\dim U = 4$ then from theorem 4.6.4 we have $U = U_2$ but $(x-2)(x-5) \in U_2, (x-2)(x-5) \notin U$, a contradiction. Thus, $\dim U \leq 3$ so dim U $= 3$ which follows $1, (x-2)(x-5)(x-6), (x-2)(x-5)(x-6)x$ is a basis of U.

   b) Add x, x^2 to the list and we get a linearly independent list with length of 5 so the new list is a basis of $\mathcal{P}_4(\mathbf{F})$.

   c) According to theorem 4.3.4 we find $W = \text{span }(x, x^2)$.

8. a) With $U = \{p \in \mathcal{P}_4(\mathbf{F}) \ : \ \int_{-1}^{1} p = 0\}$ then the list $2x, 3x^2-1, 4x^3, 5x^4-3x^2$ is a basis of U. Indeed, this list is linearly independent and from dim U ≤ 4 because $x^2 \notin U, x^2 \in \mathcal{P}_4(\mathbf{F})$, we follow $\dim U = 4$ which means the list is a basis of U according to theorem 4.6.4.

   An interesting remark is that if $V = \{p \in \mathcal{P}_5(\mathbf{F}) \ : \ p(-1) = p(1), p(0) = 0\}$ then for a basis of V, a list of antiderivative of all polynomials in basis of V is a basis of U.

   b) Add 1 to the list we get a basis of $\mathcal{P}_4(\mathbf{F})$.

   c) Similarly, $W = \text{span }(1)$.

9. If $w \notin \text{span }(v_1, \ldots, v_m)$ then $w+v_1, \ldots, w+v_m$ is linearly independent in V so $\dim \text{span}(w+v_1, \ldots, w+v_m) = m$. If $w \in \text{span }(v_1, \ldots, v_m)$ then there w can be written uniquely as $w = \sum_{i=1}^m a_i v_i$. WLOG, say the biggest i so $a_i \neq 0$ is j $(1 \leq j \leq m)$. Consider $U = \text{span}(w+v_1, \ldots, w+v_{j-1}, w+v_{j+1}, \ldots, w+v_m)$. We see that the list $w+v_1, \ldots, w+v_{j-1}, w+v_{j+1}, \ldots, w+v_m$ is linearly independent because otherwise

$$b_1(w+v_1) + \ldots + b_{j-1}(w+v_{j-1}) + b_{j+1}(w+v_{j+1}) + \ldots + b_m(w+v_m) = 0$$

is true for $b_i \in \mathbf{F}$ not all 0. However, this follows $w = c_1v_1 + \ldots + c_{j-1}v_{j-1} + c_{j+1}v_j + \ldots + c_mv_m$, a contradiction since $a_j = 0$ in this representation of w. Thus, $w+v_1, \ldots, w+v_{j-1}, w+v_{j+1}, \ldots, w+v_m$ is a basis of U so $\dim U = m-1$. On the other hand, U is a subspace of $\text{span}(w+v_1, \ldots, w+v_m)$ so

$$\dim \text{span}(w+v_1, \ldots, w+v_m) \geq \dim U = m-1.$$

10. List $p_0, p_1, \ldots, p_m$ is linearly independent. Indeed, consider $P(x) = \sum_{i=1}^{m} a_i p_i$ for $a_i \in \mathbf{F}$. We have $[x^m]P = [x^m]a_m p_m$ so $P(x) = 0$ for all $x \in \mathbf{R}$ iff $[x^m]P = 0$ or $[x^m]a_m p_m = 0$ which follows $a_m = 0$. This leads to $[x^{m-1}]P = [x^{m-1}]a_{m-1}p_{m-1}$ and we keep going until $a_i = 0$ for all $0 \leq i \leq m$. Thus, the list is linearly independent with length of $m+1 = \dim \mathcal{P}_m(\mathbf{F})$ so from theorem 4.6.4 we follow $p_0, \ldots, p_m$ is a basis of $\mathcal{P}_m(\mathbf{F})$.

11. From theorem 4.6.6 we follow $\dim (U \cap W) = 0$, i.e. $U \cap W = \{0\}$. This follows $\mathbf{R}^8 = U \oplus W$.

12. We have $\dim U = \dim W = 5$ and if $U \cap W = \{0\}$ then $\dim (U \cap W) = 0$. Hence, from theorem 4.6.6 we find $\dim (U + W) = 10$. But $U + W$ is a subspace of $\mathbf{R}^9$ which has dimension of 9, a contradiction from theorem 4.6.3. Thus $U \cap W \neq \{0\}$.

13. Since $\dim U = \dim W = 4$ and $\dim (U + W) \leq \dim \mathbf{R}^6 = 6$ so $\dim (U \cap W) \geq 2$. This follows there exists two vectors u, w in $U \cap W$ so neither of these vector is a scalar multiple of the other.

14. $\dim (U_1 + \ldots, +U_m) \leq \dim U_1 + \ldots + \dim U_m$ is true from theorem 4.6.6.

15. Since $\dim V = n$ so there exists a basis $v_1, \ldots, v_n$ of V. Pick $U_i = \operatorname{span}(v_i)$ then of course $\dim U_i = 1$ and $V = U_1 \oplus U_2 \oplus \ldots \oplus U_n$.

16. It's not hard to prove $U_1 + \ldots + U_m$ is finite-dimensional and note that let $v_{i,1}, v_{i,2}, \ldots, v_{i,b_i}$ be the basis of U_i then $\dim U_i = b_i$ and

$$v_{1,1}, \ldots, v_{1,b_1}, \ldots, v_{m,1}, \ldots, v_{m,b_m}$$

spans $U_1 + \ldots + U_m$ and is linearly independent since $U_1 + \ldots + U_m$ is a direct sum. Thus, $\dim U_1 \oplus U_2 \oplus \ldots \oplus U_m = b_1 + \ldots + b_m = \dim U_1 + \ldots + \dim U_m$.

17. This is false, take three lines in $\mathbf{R}^2$ as U_1, U_2, U_3 then $LHS = 2$ and $RHS = 3$.

5. Chapter 3: Linear Maps

5.1. 3.A The vector space of linear maps

Set of all linear maps from V to W is denoted as $\mathcal{L}(V, W)$.

theo_14:3A:3.5 **Theorem 5.1.1** (3.5 Linear maps and basis of domain) Suppose $v_1, \ldots, v_n$ is a basis of V and $w_1, \ldots, w_n \in W$. There exists a unique linear map $T : V \to W$ such that $Tv_j = w_j$ for each $j = 1, \ldots, n$.

prop:2 **Proposition 5.1.2** (Linear maps take 0 to 0) Suppose T is a linear map from V to W. Then $T(0) = 0$.

5.2. Exercises 3A

1. We have $T(0,0,0) = (0+b, 0)$ but from proposition 5.1.2 then $T(0,0,0) = (0,0)$ so $b = 0$. On the other hand, we have $T(1,1,1) = (1, 6+c)$ so $T(2,2,2) = 2T(1,1,1) = (2, 12+2c)$ but $T(2,2,2) = (2, 12+8c)$. Thus, $c = 0$. If $b = c = 0$ then it's obvious that T is linear.

2. Consider $p = 1 \in \mathcal{P}(\mathbf{R})$ then $T(1) = \left(3+b, \frac{15}{4} + c\sin 1\right)$ so $T(2) = 2T(1) = \left(6+2b, \frac{15}{2} + 2c\sin 1\right)$. However, $T(2) = \left(6+4b, \frac{15}{2} + c\sin 2\right)$. This follows $b = 0, c = 0$. The reverse is not hard.

3. Let $T\left(\underbrace{0, \ldots, 0}_{i-1}, 1, 0, \ldots, 0\right) = (A_{1,i}, A_{2,i}, \ldots, A_{m,i})$. Then, for any $(x_1, \ldots, x_n) \in \mathbf{F}^n$ we have
$$\begin{aligned} T(x_1, \ldots, x_n) &= \sum_{i=1}^n x_i \cdot T\left(\underbrace{0, \ldots, 0}_{i-1}, 1, 0, \ldots, 0\right), \\ &= \sum_{i=1}^n x_i \cdot (A_{1,i}, A_{2,i}, \ldots, A_{m,i}), \\ &= (A_{1,1}x_1 + \ldots + A_{1,n}x_n, \ldots, A_{m,1}x_1 + \ldots + A_{m,n}x_n). \end{aligned}$$

4. Assume the contrary that $v_1, \ldots, v_m$ is linearly dependent. Hence, there exists $a_i \in \mathbf{R}$ not all 0 so that $\sum_{i=1}^n a_i v_i = 0$ in V. Hence,
$$T\left(\sum_{i=1}^n a_i v_i\right) = \sum_{i=1}^n a_i T v_i = 0.$$
This follows, $Tv_1, \ldots, Tv_m$ is linearly dependent, a contradiction. Thus, $v_1, \ldots, v_m$ is linearly independent in V.

5. Yes, $\mathcal{L}(V, W)$ is a vector space.

6. Prove 3.9, page 56: Associativity: $(T_1T_2)T_3 = T_1(T_2T_3)$. Let $T_1 \in \mathcal{L}(U,V), T_2 \in \mathcal{L}(V,W), T_3 \in \mathcal{L}(W,Z)$. Hence,
$$(T_1T_2)T_3(u) = (T_1T_2)(T_3(u)) = T_1(T_2(T_3(u))) = T_1(T_2T_3).$$
7. Say we have for some $v \in V$ then $Tv = w \in V$ but since $\dim V = 1$ so $w = \lambda v$. Furthermore, according to theorem 5.1.1 there exists a unique linear map $T \in \mathcal{L}(V,V)$ so $Tv = w$ which is $Tv = \lambda v$.
8. Let $\varphi(x,y) = \sqrt[3]{xy^2}$ then $\varphi(ax,ay) = a\sqrt[3]{xy^2} = a\varphi(x,y)$ but $\varphi((x,y)+(z,t)) = \varphi(x+z, y+t) = \sqrt[3]{(x+z)(y+t)^2} \neq \varphi(x,y) + \varphi(z,t)$. Thus, φ is not linear map.
9. Let $\varphi(x+iy) = y+ix$ then $\varphi((x+iy)+(z+it)) = y+t+i(z+x) = \varphi(x+iy)+\varphi(z+it)$. But $\varphi(i(x+iy)) = \varphi(ix-y) = -iy+x \neq i\varphi(x+iy) = iy-x$. Thus, φ is not linear.
10. Consider $v_1 \in U, v_2 \in V \setminus U$ and $v_1, v_2, Sv_1 \neq 0$. Hence, $T(v_1+v_2) = 0$ since $v_1 + v_2 \notin U$ but $Tv_1 + Tv_2 = Sv_1 \neq 0$. Thus, $T(v_1+v_2) \neq Tv_1 + Tv_2$. Thus, T is not a linear map on V.
11. Consider $S \in \mathcal{L}(U,W)$ and let $v_1, \dots, v_m$ be basis of U then $Sv_i = w_i$. According to theorem 4.3.3, there exists $u_1, \dots, u_l$ so $v_1, \dots, v_m, u_1, \dots, u_l$ is a basis of V. Pick arbitrary l vectors $\ell_1, \dots, \ell_l$ in W. According to theorem 5.1.1, there exists a linear map $T : V \to W$ so $Tv_i = w_i = Sv_i, Tu_i = \ell_i$.
12. Since V is finite-dimensional so there exists a basis $v_1, \dots, v_m$ of V. Since W is infinite-dimensional so according to Exercise 14 (2A), there exists sequence of vectors $w_1, w_2, \dots$ so that $w_1, w_2, \dots, w_k$ is linearly independent for any $k \in \mathbf{N}$. In order to show that $\mathcal{L}(V,W)$ is infinite-dimensional, we will prove the existence of seuquences $T_1, T_2, \dots$ where $T_i \in \mathcal{L}(V,W)$ and $T_1, \dots, T_k$ is linearly independent for any $k \in \mathbf{N}$.

   Indeed, according to theorem 5.1.1, there exists $T_i : V \to W$ so that $T_i(v_j) = w_{(i-1)m+j}$ for all $1 \le j \le m$. Thus, for $v = \sum_{j=1}^m \alpha_j v_j$ then $T_i(v) = \sum_{j=1}^m \alpha_j w_{(i-1)m+j}$. Thus, for any $k \in \mathbf{N}$ and for all $v \in V, v \neq 0, v = \sum_{j=1}^m \alpha_j v_j$ if
$$\sum_{i=1}^k \beta_i T_i(v) = 0v \iff \sum_{i=1}^k \sum_{j=1}^m \beta_i \alpha_j w_{(i-1)m+j} = 0v.$$
   Note that $w_1, \dots, w_{km}$ is linearly independent so the above equation follows that $\beta_i\alpha_j = 0$ for any $1 \le i \le k, 1 \le j \le m$. However, since $v \neq 0$ so there must exists an $\alpha_l \neq 0$ $(1 \le l \le m)$. This deduces $\beta_i = 0$ for all $1 \le i \le k$, i.e. $T_1, \dots, T_k$ is linearly independent. And this is true for any $k \in \mathbf{N}$ so $\mathcal{L}(V,W)$ is infinite-dimensional.
13. Since $v_1, \dots, v_m$ is linear dependent so there exists $v_i = \sum_{j \neq i} \alpha_j v_j$ with $\alpha_j \in \mathbf{F}$. Hence, pick arbitrary w_j $(1 \le j \le m, j \neq i)$ and then pick w_i so $w_i \neq \sum_{j \neq i} \alpha_j w_j = Tv_i$. Done.
14. Let $v_1, \dots, v_m$ $(m \ge 2)$ be basis of V then for $i \ge 3$, let $Sv_i = Tv_i = v_i$ but $Sv_1 = Tv_2 = -v_1, Sv_2 = Tv_1 = v_2$. Hence, $ST(v_1) = Sv_2 = v_2$ but $TS(v_1) = T(-v_1) = -v_2$. Thus, $ST \neq TS$.

5.3. 3.B Null Spaces and Ranges

Proposition 5.3.1
(3.14) Suppose $T \in \mathcal{L}(V, W)$ then null T is a subspace of V.

Proof. According to the definition of null space if $v \in$ null T then $v \in V$. If $v_1, v_2 \in$ null T then $Tv_1 = Tv_2 = 0$ so $T(v_1 + v_2) = Tv_1 + Tv_2 = 0$. Hence, $v_1 + v_2 \in$ null T. We also have $T(\lambda v_1) = \lambda T v_1 = 0$ so $\lambda v_1 \in$ null T for any $\lambda \in \mathbf{F}$. □

Proposition 5.3.2
(3.16) Let $T \in \mathcal{L}(V, W)$ then T is injective if and only if null $T = \{0\}$.

Proof. If T is injective. Assume the contrary that there exists $v \neq 0, v \in$ null T. then for any $u \neq v, u \in V$ then $T(u + v) = Tu + Tv = Tu$, a contradiction since T is injective.

Reversely, if null T$T = \{0\}$. For any $u, v \in V$ so $Tu = Tv$ then $T(u - v) = 0$. Thus, $u = v$. This follows T is injective. □

Proposition 5.3.3
(3.19) If $T \in \mathcal{L}(V, W)$ then range T is a subspace of W.

Proof. We have $T(0) = 0$ so $0 \in$ range T. For any $v_1, v_2 \in$ range T then there exists $u_1, u_2 \in V$ so $Tu_1 = v_1, Tu_2 = v_2$ so $T(u_1 + u_2) = v_1 + v_2$. Hence, $v_1 + v_2 \in$ range T. We also have $\lambda v_1 = \lambda T u_1 = T(\lambda u_1)$ so $\lambda v_1 \in$ range T for any $\lambda \in \mathbf{F}$. Thus, range T is a subspace of W. □

Definition 5.3.4. A function $T : V \to W$ is called surjective if its range equal W.

theo_15:3B:3.22 **Theorem 5.3.5** (3.22, Fundamental Theorem of Linear Maps) Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Then range T is finite-dimensional and

$$\dim V = \dim \text{null } T + \dim \text{range } T.$$

Proof. Let $v_1, \ldots, v_m$ be the basis of V then for any $v \in V, v = \sum_{i=1}^m \alpha_i v_i$ then $f(v) \in$ range T and $f(v) = \sum_{i=1}^m \alpha_i f(v_i)$. This follows range $T =$ span $(f(v_1), \ldots, f(v_m))$ or range T is finite-dimensional.

Let $u_1, \ldots, u_{m-k}$ be basis of null T. Hence, according to theorem 4.3.3, linearly independent list $u_1, \ldots, u_{m-k}$ can be extented to a basis $u_1, \ldots, u_{m-k}, w_1, \ldots, w_k$ of V. It suffices to prove that dim range $T = k$.

Indeed, let $W = \text{span}(w_1, \ldots, w_k)$. Note that for any $u, v \in W, v \neq u$ then $Tv \neq Tu$, otherwise $T(u - v) = 0$ which means $, u - v \neq 0, u - v \in$ null T but $u - v \in W$, a contradiction since $u_1, \ldots, u_{m-k}, w_1, \ldots, w_k$ is linearly independent.

We prove that $(Tw_1, Tw_2, \ldots, Tw_k)$ is a basis of range T. From the above argument, we find $Tw_1, \ldots, Tw_k$ is linearly independent. For any $v \in$ range T then there exists $u \in V$ so $Tu = v$.

According to theorem 4.3.4, $V = \text{null } T \oplus W$ so for any $u \in V$ there uniquely exists $u_1 \in \text{null } T, w \in W$ so $u = u_1 + w$. Hence, $Tu = Tu_1 + Tw = Tw$ or $v = Tu = Tw = \sum_{i=1}^{k} \alpha_k T w_k$ for any $v \in \text{range } T$. Thus, $Tw_1, \ldots, Tw_k$ spans range T, which follows $\dim \text{range } T = k$, as desired. ☐

theo_16:3B:3.23 **Theorem 5.3.6** (3.23) Suppose V and W are finite-dimensional vector spaces such that $\dim V > \dim W$. Then no linear map from V to W is injective.

Proof. Asusme the contrary, there is injective linear map $T \in \mathcal{L}(V, W)$. Hence, $\text{null } T = \{0\}$ or $\dim \text{null } T = 0$. Thus, from theorem 5.3.5, $\dim V = \dim \text{range } T \le \dim W$ since range T is a subspace of W, a contradiction. Thus, there doesn't exist injective linear map from V to W. ☐

theo_16:3B:3.24 **Theorem 5.3.7** (3.24) Suppose V and W are finite-dimensional vector spaces such that $\dim V < \dim W$. Then no linear map from V to W is surjective.

Proof. For any $T \in \mathcal{L}(V, W)$, we have $\dim \text{range } T = \dim V - \dim \text{null } T < \dim W$, which means range $T \neq W$, so T is not surjective. ☐

Theorem 5.3.8

(3.26) A homogeneous system of linear equations with more variables than equations has nonzero solutions.

Proof. Using the notation from the textbook. This means $n > m$, or $\dim \mathbf{F}^n > \dim \mathbf{F}^m$. This follows no linear map from $\mathbf{F}^n$ to $\mathbf{F}^m$ is injective, i.e. $\dim \text{null} T(x_1, \ldots, x_n) > 0$ for any $T \in \mathcal{L}(V, W)$. This means the system of linear equations has nonzero solutions. ☐

Theorem 5.3.9

(3.29) An inhomogeneous system of linear equations with more equations than variables has no solution for some choice of constant terms.

Proof. We have $\dim \mathbf{F}^n < \dim \mathbf{F}^m$ so no linear map from $\mathbf{F}^n$ to $\mathbf{F}^m$ is surjective. Hence, for any $T \in \mathcal{L}(V, W)$ then range $T \neq \mathbf{F}^m$. This follows there exists $(c_1, \ldots, c_m) \in \mathbf{F}^m$ but $(c_1, \ldots, c_m) \notin \text{range } T$. Thus, the system of linear equations with constant terms $c_1, \ldots, c_m$ has no solution. ☐

5.4. Exercises 3B

5.4.1. A way to construct (not) injective, surjective linear map

That is based on theorem 5.1.1. T is injective if basis V is mapped one-to-one with basis of W (which is also we need $\dim V \le \dim W$ for existence of injective linear map from V to W). In other words, if $v_1, \ldots, v_m$ is basis of V then $T \in \mathcal{L}(V, W)$ is injective if and only if $Tv_1, \ldots, Tv_m$ are linearly independent in W. T is not injective otherwise.

Similarly, T is surjective if and only if $Tv_1, \ldots, Tv_m$ spans W.

Some examples are exercises 7, 8, 9, 10.

Many exercises below use theorem 5.1.1!

5.4.2. Exercises

1. If $T \in \mathcal{L}(V, W)$ then from theorem 5.3.5, $\dim V = 5$. We can let $V = \mathbf{F}^5$ and since range T is a subspace of W so $2 = \dim \operatorname{range} T \le \dim W$. Hence, we can let $W = \mathbf{F}^2$. Hence, we can write $T(x, y, z, t, w) = (x + y, t + w)$. With this, null $T = \{(x, y, z, t, w) \in \mathbf{F}^5 : x + y = t + w = 0\}$. Hence, basis of null T is $(1, -1, 0, 0, 0), (0, 0, 1, 0, 0), (0, 0, -1, 1)$. And range $T = \mathbf{F}^2$.

2. Since range $S \subset$ null T so $(TS)(v) = T(Sv) = 0$ because $Sv \in$ range S so $Sv \in$ null T. Thus,
$$(ST)^2(v) = (ST)((ST)v) = ST(Su) = S(TSu) = S(0) = 0.$$

3. (a) If $v_1, \ldots, v_m$ spans V then range $T = V$ or T is surjective.
(b) If $v_1, \ldots, v_m$ are linearly independent then $T(z_1, \ldots, z_m) = 0$ if and only if $(z_1, \ldots, z_m) = 0$, i.e. null $T = \{0\}$ or T is injective.

4. Let $T_1(x, y, z, t, w) = (x + y, 0, 0, t + z + w)$ then range $T_1 = \{(x, 0, 0, y) \in \mathbf{R}^4\}$ so $\dim \operatorname{range} T_1 = 2$ so $\dim \operatorname{null} T_1 = 3$. Let $T_2(x, y, z, t, w) = (x + y, 0, t, 0)$ then $(T_1 + T_2)(x, y, z, t, w) = (2x + 2y, 0, t, t + z + w)$ so $\dim \operatorname{range} (T_1 + T_2) = 3$. This follows $\dim \operatorname{null} (T_1 + T_2) = 2$ or $T_1 + T_2 \notin S = \{T \in \mathcal{L}(\mathbf{R}^5, \mathbf{R}^4) : \dim \operatorname{null} T > 2\}$. Hence, S is not a subspace of $\mathcal{L}(\mathbf{R^5}, \mathbf{R}^4)$.

5. We contruct such linear map based on proof of theorem 5.3.5. Note that basis of $\mathbf{R}^4$ is $(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)$ so if we choose $(1, 0, 0, 0), (0, 1, 0, 0)$ to be basis of null T then we need $T(0, 0, 1, 0), T(0, 0, 0, 1)$ to be basis of range T. Therefore, null $T = \{(x, y, 0, 0) \in \mathbf{R}^4\} =$ range T so we can set $T(0, 0, 1, 0) = (1, 0, 0, 0), T(0, 0, 0, 1) = (0, 1, 0, 0)$. Hence, T can be $T(x, y, z, t) = (z, t, 0, 0)$.

6. Because that will mean $\dim \operatorname{null} T = \dim \operatorname{range} T$ so $5 = \dim \mathbf{R}^5 = 2 \cdot \dim \operatorname{null} T$, a contradiction.

exer:3B:7

7. Let $v_1, \ldots, v_m$ be basis of V and $w_1, \ldots, w_n$ be basis of W with $n \ge m \ge 2$. Construct T so $Tv_i = w_i$ for $1 \le i \le m - 1$ but $Tv_m = 2w_{m-1}$. Next construct S so $Sv_i = w_i$ for $1 \le i \le m - 2$ but $Sv_{m-1} = 2w_m$ and $Sv_m = w_m$. It's obvious that $T, S \in R = \{T \in \mathcal{L}(V, W) : T \text{ is not injective}\}$.

   However, $T + S$ is injective. Indeed, we have $(T + S)(v_i) = w_i$ for $1 \le i \le m - 2$ and $(T + S)v_{m-1} = w_{m-1} + 2w_m, (T + S)v_m = 2w_{m-1} + w_m$. Note that $w_1, \ldots, w_{m-2}, w_{m-1} + 2w_m, 2w_{m-1} + w_m$ are linearly independent so dim null $(T + S) = \{0\}$ or $T + S$ is injective. Thus, R is not subspace of $\mathcal{L}(V, W)$.

exer:3B:8 8. Notation the same with exercise 7 but this time $m \geq n \geq 2$. Construct T so $Tv_i = w_i$ for $1 \leq i \leq n-2$, $Tv_j = jw_{n-1}$ for $n-1 \leq j \leq m$. Construct S so $Sv_i = w_i$ for $1 \leq i \leq n-2$, $Tv_j = (m+1-j)w_n$ for $n-1 \leq j \leq m$. It's obvious that $T, S \in R = \{T \in \mathcal{L}(V,W) : T \text{ is not surjective}\}$.

   However, $T+S$ is surjective since $Tv_1, Tv_2, \ldots, Tv_n$ is $w_1, \ldots, w_{n-2}, (n-1)w_{n-1} + (m-n+2)w_n, nw_{n-1} + (m-n+1)w_n, \ldots$ which spans W. Thus, $S+T$ is surjective. Thus, R is not subspace of $\mathcal{L}(V,W)$.

exer:3B:9 9. T is injective so for any $v \in \text{span}\,(v_1, \ldots, w_m)$ then $Tv = 0$ if and only if $v = 0$, i.e. $\sum_{i=1}^m \alpha_i Tv_i = 0$ if and only if $\alpha_i = 0$ for all $1 \leq i \leq m$, i.e. $Tv_1, \ldots, Tv_m$ is linearly independent in W.

exer:3B:10 10. For any $u \in \text{range}\, T$ there exists $v \in V$ or $v = \sum_{i=1}^m \alpha_i v_i$ so $Tv = u$ or $u = \sum_{i=1}^m \alpha_i Tv_i$. Thus, any $u \in \text{range}\, T$ can be expressed as linear combination of $Tv_1, \ldots, Tv_m$ which follows $Tv_1, \ldots, Tv_m$ spans range T.

exer:3B:11 11. If we have $S_1(S_2 \ldots S_n u) = S_1(S_2 \ldots S_n v)$ then we find $S_2 \ldots S_n u = S_2 \ldots S_n v$ since S_1 is injective linear map. Similarly, we from this we deduce $u = v$. This follows $S_1 S_2 \ldots S_n$ is injective.

exer:3B:12 12. Since null T is a subspace of finite-dimensional vector space V so according to theorem 4.3.3, a basis $v_1, \ldots, v_m$ of null T can be extended to a basis of V, i.e. $v_1, \ldots v_m, u_1, \ldots, u_n$. Hence, let $U = \text{span}\,(u_1, \ldots, u_n)$ then we have $U \cap \text{null}\, T = \{0\}$. On the other hand, any $v \in V$ there exists $u \in U, z \in \text{null}\, T$ so $v = u + z$ so $Tv = T(u+z) = Tu$. This follows range $T = \{Tv : v \in V\} = \{Tu : u \in U\}$.

exer:3B:13 13. A basis for null T is $(5,1,0,0), (0,0,7,1)$ so dim null $T = 2$. Combining with theorem 5.3.5 we find dim range $T = 2 = \dim W$ and since range T is a subspace of W so range $T = W$ according to theorem 4.6.4. Thus, T is surjective.

exer:3B:14 14. Similarly, from theorem 5.3.5 to find dim range $T = 5 = \dim \mathbf{R}^5$ so T is surjective.

exer:3B:15 15. Assume the contrary, there exists such linear map T then dim null $T = 2$. On the other hand, dim range $T \leq \dim \mathbf{R}^2 = 2$ so dim null T + dim range $T \leq 4$, a contradiction to theorem 5.3.5.

exer:3B:16 16. Assume the contrary that V is infinite-dimensional then according to exercise 14 there exists a sequence $v_1, v_2, \ldots$ so $v_1, \ldots, v_m$ is linearly independent for every positive integer m. WLOG, say that $v_1, \ldots, v_n$ is basis of null T.

   Since range T of a linear map on V is finite-dimensional so there exists $u_1, \ldots, u_k \in V$ so $Tu_1, \ldots, Tu_k$ is a basis of range T. Note that there exists a $M > n$ so all u_j $(1 \leq j \leq k)$ can be represented as linear combination of v_i for $i < M$. Hence, since $Tv_M \in \text{range}\, T$ so

$$Tv_M = \sum_{i=1}^{\ell} \alpha_i Tu_i = \sum_{i=1}^{M-1} \beta_i Tv_i.$$

This follows $v_M - \sum_{i=1}^{M-1} \beta_i v_i \in \text{null } T$ or $v_M = \sum_{i=1}^{M-1} \beta_i v_i + \sum_{j=1}^{n} \gamma_j v_j$, a contradiction since $v_1, \dots, v_M$ is linearly independent. Thus, V is finite-dimensional.

exer:3B:17 17. According to theorem 5.3.6, if there exists an injective linear map from V to W then $\dim V \le \dim W$. Conversely, if $\dim V \le \dim W$, let $v_1, \dots, v_m$ be basis of V and $w_1, \dots, w_n$ $(n \ge m)$ be basis of W then according to theorem 5.1.1, there exists a linear map $T : V \to W$ such that $Tv_j = w_j$ for $1 \le j \le m$. With this, we will find that $\text{null } T = \{0\}$ so T is injective.

exer:3B:18 18. According to theorem 5.3.7, if there exists a surjective map from V onto W then $\dim V \ge \dim W$. Conversely, if $\dim V \ge \dim W$, let $v_1, \dots, v_m$ be basis of V and $w_1, \dots, w_n$ $(n \le m)$ be basis of W then according to theorem 5.1.1, there exists a linear map $T : V \to W$ such that $Tv_j = w_j$ for $1 \le j \le n$. With this, we find $\text{range } T = W$ so T is surjective.

exer:3B:19 19. If there exists such linear map T then according to theorem 5.3.5, we have $\dim \text{null } T = \dim U = \dim V - \dim \text{range } T \ge \dim V - \dim W$. Conversely, if $\dim U \ge \dim V - \dim W$, let $u_1, u_2, \dots, u_m, v_1, \dots, v_n$ be basis of V where $u_1, \dots, u_m$ be basis of U and let $w_1, \dots, w_k$ be basis of W then $k \ge n$. According to theorem 5.1.1, there exists a linear map $T \in \mathcal{L}(V, W)$ so $Tu_i = 0, Tv_j = w_j$ for $1 \le j \le n, 1 \le i \le m$. With $k \ge n$, we can show that $\text{null } T = U$.

exer:3B:20 20. If there exists such linear map S then when $Tu = Tv$ then $STu = STv$ or $u = v$. Thus, T is injective. Conversely, if T is injective, let $Tv_1, \dots, Tv_m$ be a basis of range T. Hence, according to 5.1.1, there exists a linear map $S \in \mathcal{L}(W, V)$ so $S(Tv_i) = v_i$ and $S(w_j) = 0$ for arbitrary $u_j \in V$ where $Tv_1, \dots, Tv_m, w_1, \dots, w_n$ is basis of W.

For any $v \in V$ then $Tv = \sum_{i=1}^{m} \alpha_i Tv_i$. Since T is injective so $\text{null } T = 0$ so $v = \sum_{i=1}^{m} \alpha_i v_i$. Hence,

$$STv = S\left(\sum_{i=1}^{m} \alpha_i Tv_i\right) = \sum_{i=1}^{m} \alpha_i STv_i = \sum_{i=1}^{m} \alpha_i v_i = v.$$

exer:3B:21 21. If there exists such linear map S then any $w \in W$ we have $T(Sw) = w \in W$ so $w \in \text{range } T$. Since range T is a subspace of W, this follows range $T = W$ or T is surjective.

Conversely, if T is surjective then range $T = W$. There exists $v_1, \dots, v_m \in V$ so $Tv_1, \dots, Tv_m$ is basis of W. Hence, according to theorem 5.1.1, there exists a linear map $S \in \mathcal{L}(W, V)$ so $S(Tv_i) = v_i$ for all $1 \le i \le m$. This follows, for any $w \in W, w = \sum_{i=1}^{m} \alpha_i Tv_i$ then

$$TSw = T\left(\sum_{i=1}^{m} \alpha_i S(Tv_i)\right) = T\left(\sum_{i=1}^{m} \alpha_i v_i\right) = \sum_{i=1}^{m} \alpha_i Tv_i = w.$$

exer:3B:22 22. For any $v \in \text{null } ST$ then $Tv \in \text{null } S$. Hence, we define linear map $T' \in \mathcal{L}(\text{null } ST, \text{null } S)$ so that for any $v \in \text{null } ST$ then $T'v = Tv$. Note that since null ST is a subspace of V so null T' is a subspace of null T so $\dim \text{null } T' \le \dim \text{null } T$. According to theorem 5.3.5, we have

$$\dim \text{null } ST = \dim \text{null } T' + \dim \text{range } T' \le \dim \text{null } T + \dim \text{null } S.$$

exer:3B:23 23. Note that range ST is a subspace of range S so dim range $ST \leq$ dim range S. We can write range $ST = \{S(Tv) : v \in U\} = \{Sv : v \in \text{range } T\}$. Let $u_1, \dots, u_m$ be basis of range T then $Su_1, \dots, Su_m$ will span range ST so dim range $ST \leq$ dim range T. Thus, dim range $ST \leq \min\{\text{dim range } S, \text{dim range } T\}$.

exer:3B:24 24. (note that V is not necessarily finite-dimensional so we can't actually set a basis for V) If there exists such S, then for any $v \in \text{null } T_1$ then $ST_1v = S(0) = 0 = T_2v$ or $v \in \text{null } T_2$. Thus, null $T_1 \subset$ null T_2.

Conversely, if null $T_1 \subset$ null T_2, let $T_1v_1, \dots, T_1v_m, w_1, \dots, w_n$ be basis of W where $Tv_1, \dots, Tv_m$ is basis of range T. Let S be linear map so $S(Tv_i) = T_2v_i$ and $Sw_j = 0$ for $1 \leq i \leq m, 1 \leq j \leq n$.

We prove that $T_2 = ST_1$. Indeed, for any $v \in V$, we have $Tv \in \text{range } T$ so $T_1v = \sum_{i=1}^m \alpha_i T_1 v_i$ so $v - \sum_{i=1}^m \alpha_i v_i \in \text{null } T_1 \subset \text{null } T_2$. Hence,

$$T_2(v) = T_2\left(\sum_{i=1}^m \alpha_i v_i\right) = \sum_{i=1}^m \alpha_i T_2 v_i = \sum_{i=1}^m \alpha_i ST_1 v_i = ST_1\left(\sum_{i=1}^m \alpha_i T_1 v_i\right) = S(T_1v).$$

exer:3B:25 25. If there exists such S then for all $v \in V$ we have $T_1v = T_2Sv \in \text{range } T_2$ which follows range $T_1 \subset$ range T_2. Conversely, for $v_1, \dots, v_m$ as basis of V, since range $T_1 \subset$ range T_2 there exists u_i so $T_1v_i = T_2u_i$ for all $1 \leq i \leq m$. According to theorem 5.1.1, there exists a linear map $S \in \mathcal{L}(V, V)$ so $Sv_i = u_i$ for all $1 \leq i \leq m$. This follows $T_2Sv_i = T_2u_i = T_1v_i$ for all $1 \leq i \leq m$, which leads to $T_2Sv = T_1v$ for all $v \in V$.

exer:3B:26 26. Let $Dx^i = p_i$ for all $i \geq 1$ then $\deg p_i = i - 1$ so every polynomial can be represented as linear combination of p_i. For any $p' \in \mathcal{P}(\mathbf{R})$, $p' = \sum_{i=1}^m \alpha_i p^i$ then $D(\sum_{i=1}^m \alpha_i x^i) = p'$. Hence, $p' \in \text{range } D$. This follows range $D = \mathcal{P}(\mathbf{R})$ or D is surjective.

exer:3B:27 27. Consider a linear map $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathcal{P}(\mathbf{R}))$ so $Dp = 5p'' + 3p'$ then according to exercise 26, D is surjective so for any $p \in \mathcal{P}(\mathbf{R})$ then there exists $q \in \mathcal{P}(\mathbb{R})$ so $Dq = p$ or $5q'' + 3q' = p$.

exer:3B:28 28. Let $\varphi_i(v)$ be scalar multiple of w_i in representation of Tv. We prove that φ_i is linear. Indeed, since T is linear so $T(u+v) = Tu + Tv$ or $\sum_{i=1}^m \varphi_i(u+v)w_i = \sum_{i=1}^m (\varphi_i u + \varphi_i v)w_i$. Since $w_1, \dots, w_m$ is linearly independent so $\varphi_i(u+v) = \varphi_i(u) + \varphi_i(v)$. Similarly, $\varphi_i(\lambda v) = \lambda \varphi_i v$ for any $v \in V$.

exer:3B:29 29. It's obvious that null $\varphi \oplus \{au : a \in \mathbf{F}\}$ is subset of V. On the other hand, for any $v \in V$ then $\varphi(v) = a$. We know $\varphi(u) \neq 0$ so $x = v - \frac{a}{\varphi u}u \in V$. Hence, we have

$$a = \varphi(v) = \varphi\left(x + \frac{a}{\varphi u}u\right) = a + \varphi(x).$$

Thus, $\varphi x = 0$ or $x \in \text{null } \varphi$. Hence, any $v \in V$ can be represented uniquely as $v = w + au$ where $w \in \text{null } \varphi$ and $a \in \mathbf{F}$. This follows null $\varphi \oplus \{au : a \in \mathbf{F}\} = V$.

exer:3B:30 30. According to exercise 29, if there is $u \in V$ not in $\operatorname{null} \varphi_1 = \operatorname{null} \varphi_2$ then let $c = \varphi_1(u)/\varphi_2(u)$. For any $v \in V$, v can be represented as $v = z + au$, $(a \in \mathbf{F})$ where $z \in \operatorname{null} \varphi_1$. Thus, $\varphi_1(v) = \varphi_1(au) = a\varphi_1(u) = c\varphi_2(v)$.

exer:3B:31 31. Null space of T_1, T_2 is $\{(x_1, x_2, x_3, 0, 0) : x_1, x_2, x_3 \in \mathbf{R}\}$. We have $T_1(0,0,0,1,0) = T_2(0,0,0,0,1) = (1,0)$ and $T_1(0,0,0,0,1) = T_2(0,0,0,1,0) = (0,1)$.

5.5. Exercises 3C

exer:3C:1 1. Let $v_1, \ldots, v_n$ be basis of V and $w_1, \ldots, w_m$ be basis of W where $w_1, \ldots, w_k$ $(k = \dim \operatorname{range} T \leq m)$ is basis of range T since range T is subspace of W. We know that $Tv_1, \ldots, Tv_n$ spans range T and for any $1 \leq i \leq k$ then $w_i \in \operatorname{range} T$, hence w_i can be written as linear combination of Tv_j $(1 \leq j \leq n)$. We have

$$w_i = \sum_{j=1}^{n} \alpha_j T v_j = \sum_{j=1}^{n} \alpha_j \sum_{h=1}^{m} A_{h,j} w_j.$$

In order for the above to hold, there must exist a $1 \leq h \leq m$ so $A_{h,i} \neq 0$. This is true for any $1 \leq i \leq k$. Thus, $\mathcal{M}(T)$ has at least $k = \dim \operatorname{range} T$ nonzero entries.

exer:3C:2 2. Let $v_1, \ldots, v_4$ be basis of $\mathcal{P}_3(\mathbf{R})$ and $w_1, \ldots, w_3$ be basis of $\mathcal{P}_2(\mathbf{R})$. From the matrix, since $A_{i,j} = 0$ for $i \neq j$ and $A_{i,i} = 0$ so we find $Dv_k = w_k$ for $1 \leq k \leq 3$ and $Dv_4 = 0$. So we can choose $v_i = x^i$ for $1 \leq i \leq 3$ and $v_4 = 0$ then $w_i = ix^{i-1}$ for $1 \leq i \leq 3$.

exer:3C:3 3. Let $Tv_1, \ldots, Tv_m$ be basis of range T where $v_i \in V$ then $v_1, \ldots, v_m$ is linearly independent in V, hence $u_1, \ldots, u_n$ can be added to the list to create a basis of V. This follows that $u_1, \ldots, u_n$ is basis of null space of T. Thus, with $v_1, \ldots, v_m, u_1, \ldots, u_n$ as basis of V and $Tv_1, \ldots, Tv_m$ as basis of range T, we have $A_{j,j} = 1$ for $1 \leq j \leq \dim \operatorname{range} T$.

exer:3C:4 4. Since the first column of $\mathcal{M}(T)$ (except for possibly $A_{1,1} = 1$) so we can let $w_1 = Tv_1$. It doesn't matter what should $w_2, \ldots, w_n$ specifically be as long as $w_1, \ldots, w_n$ is basis of W.

exer:3C:5 5. If there doesn't exists $v = w_1 + \sum_{i=2}^{n} \alpha_i w_i$ in range T then for any $u \in V$, since $Tv_i \in \operatorname{range} T$ so we must have $A_{1,i} = 0$ for all $1 \leq i \leq m$, or entries of the first row of $\mathcal{M}(T)$ are all 0.

If there does exist such $w = w_1 + \sum_{i=2}^{n} \alpha_i w_i$ in range T then that means there exists $v_1 \in V$ so $Tv_1 = w$. Extend v_1 to a basis $v_1, \ldots, v_m$ of V with $Tv_i = u_i$ for $i \geq 2$. Note that if $v_1, \ldots, v_m$ is basis of of V then so does $v_1, v_2 - \beta_2 v_1, v_3 - \beta_3 v_1, \ldots, v_m - \beta_m v_1$. Hence, $T(v_i - \beta_i v_1) = u_i - \beta_i w$. By choosing the right β_i to remove w_1 from u_i, we find that $v_1, v_2 - \beta_2 v_1, v_3 - \beta_3 v_1, \ldots, v_m - \beta_m v_1$ is the desired basis of V, which will gives $A_{1,k} = 0$ for all $1 \leq k \leq m$.

exer:3C:6 6. If there exists such bases of V, W then $Tv_1 = Tv_2 = \ldots = Tv_m \neq 0$. Since $v_1, v_2, \ldots, v_m$ is basis of V so $v_1, v_2 - v_1, v_3 - v_1, \ldots, v_m - v_1$ is also basis of V. Note that $T(v_i - v_1) = 0$ for all $2 \leq i \leq m$ so that means $\dim \operatorname{null} V = m - 1$ or $\dim \operatorname{range} T = 1$ according to theorem 5.3.5.

Conversely, if dim range $T = 1$ then dim null $V = m-1$ so there exists basis $v_2, \ldots, v_m$ of null V. This list can be exented to basis $v_1, \ldots, v_m$ of V. Hence, $Tv_1 = T(v_1+v_2) = \ldots = T(v_1+v_m)$ and $v_1, v_1+v_2, \ldots, v_1+v_m$ is basis of V. From this, we can choose basis of W to obtain matrix $\mathcal{M}(T)$ of entries 1.

exer:3C:7 7. Let $\mathcal{M}(T) = A, \mathcal{M}(S) = C$ then $Tv_k = A_{1,k}w_1 + \ldots + A_{m,k}w_m$ and $Sv_k = C_{1,k}w_1 + \ldots + C_{m,k}w_m$. This follows $(T+S)v_k = \sum_{i=1}^m (A_{i,k}+C_{i,k})w_i$. Thus, (i,k)th entry of $\mathcal{M}(T+S)$ is $A_{i,k}+C_{i,k}$, which also equals to (i,k)th entry of $\mathcal{M}(T)+\mathcal{M}(S)$.

exer:3C:8 8. We have $(\lambda T)v_k = \lambda(Tv_k)$.

exer:3C:9 9. Done.

exer:3C:10 10. It suffices to prove $(AC)_{j,k} = (A_{j,.}C)_{1,k}$ for any $1 \leq k \leq p$. Indeed,

$$(A_{j,.}C)_{1,k} = \sum_{i=1}^n (A_{j,.})_{1,i}C_{i,k} = \sum_{i=1}^n A_{j,i}C_{i,k} = (AC)_{j,k}.$$

exer:3C:11 11. We have

$$\begin{aligned}(a_1C_{1,.} + \cdots + a_nC_{n,.})_{1,k} &= (a_1C_{1,.})_{1,k} + \ldots + (a_nC_{n,.})_{1,k}, \\ &= \sum_{i=1}^n a_iC_{i,k} = (aC)_{i,k}.\end{aligned}$$

exer:3C:12 12. Pick $A = \begin{pmatrix} 1,0 \\ 0,0 \end{pmatrix}$ and $C = \begin{pmatrix} 1,1 \\ 0,0 \end{pmatrix}$ then $AC = \begin{pmatrix} 1,1 \\ 0,0 \end{pmatrix}$ but $CA = \begin{pmatrix} 1,0 \\ 0,0 \end{pmatrix}$.

exer:3C:13 13. Prove $A(B+C) = AB+AC$, say A is m-by-n matrix and B, C are n-by-p matrix. Hence,

$$\begin{aligned}(AB+AC)_{i,j} = (AB)_{i,j} + (AC)_{i,j} &= \sum_{h=1}^n (A_{i,h}B_{h,j} + A_{i,h}C_{h,j}) \\ &= \sum_{h=1}^n A_{i,h}(B+C)_{h,j} = (A(B+C))_{i,j}.\end{aligned}$$

exer:3C:14 14. A, B, C are m-by-n,n-by-p,p-by-q matrices then

$$\begin{aligned}((AB)C)_{i,j} &= \sum_{h=1}^p (AB)_{i,h}C_{h,j}, \\ &= \sum_{h=1}^p C_{h,j} \sum_{l=1}^n A_{i,l}B_{l,h}, \\ &= \sum_{l=1}^n A_{i,l} \sum_{h=1}^p B_{l,h}C_{h,j}, \\ &= \sum_{l=1}^n A_{i,l}(BC)_{l,j}, = (A(BC))_{i,j}.\end{aligned}$$

exer:3C:15 15. Just did in 14.

5.6. 3D: Invertibility and Isomorphic Vector Spaces

Theorem 5.6.1 (3.56)
A linear map is **invertible** or **non-singular** if and only if it is injective and surjective.

This is deduced from exercises 20 and 21 (3B).

theo:3.59:3D **Theorem 5.6.2** (3.59) Two finite-dimensional vector space over $\mathbf{F}$ are isomorphic if and only if they have same dimension.

theo:3.65:3D **Theorem 5.6.3** (3.65) Suppose $T \in \mathcal{L}(V, W)$ and $v\ inV$. Suppose $v_1, \ldots, v_n$ is basis of V and $w_1, \ldots, w_m$ is basis of W. Then

$$\mathcal{M}(Tv) = \mathcal{M}(T)\mathcal{T}(v).$$

theo:3.69:3D **Theorem 5.6.4** (3.69) Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then the following are equivalent:

(a) T is invertible.

(b) T is surjective.

(c) T is injective.

5.7. Exercises 3D

exer:3D:1 1. We have $(ST)(T^{-1}S^{-1}) = I$ and $(T^{-1}S^{-1})(ST) = I$.

exer:3D:2 2. Let $v_1, \ldots, v_n$ be basis of V. Consider noninvertible operators $T_i \in \mathcal{L}(V)$ so $T_i(v_j) = 0$ if $j \neq i$ and $T_i(v_i) = v_i$. Hence, linear map $T_1 + \ldots + T_n$ is the identity map, which is invertible. Thus, set of noninvertible operators on V is not a subspace of $\mathcal{L}(V)$.

exer:3D:3 3. If there exists such invertible operator T then if $Su = Sv$ then $Tu = Tv$ which follows $u = v$. Thus, S is injective.

   If S is injective, since U is subspace of V, basis $v_1, \ldots, v_m$ of U can be extended to basis $v_1, \ldots, v_n (m \leq n)$ of V. Since S is injective so $Sv_1, \ldots, Sv_m$ is linearly independent, so this list can be extended to basis $Sv_1, \ldots, Sv_m, u_{m+1}, \ldots, u_n$ of V. From this, we let T be a linear map so $Tv_i = Sv_i$ for $1 \leq i \leq m$ and $Tv_i = u_i$ for $m < i \leq n$. We find that T is surjective so T is invertible and $Tu = Su$ for any $u \in U$.

exer:3D:4 4. If there exists such invertible operator S then if $u \in \operatorname{null} T_1$ then $0 = T_1 u = ST_2 u$ so $T_2 u = 0$. Similarly if $u \in \operatorname{null} T_2$ then $u \in \operatorname{null} T_1$. Thus, $\operatorname{null} T_1 = \operatorname{null} T_2$.

If null T_1 = null T_2: Let $T_2v_1, T_2v_2, \ldots, T_2v_m, w_1, \ldots, w_n$ be basis of W. Note that $T_1v_1, \ldots, T_1v_m$ is linearly independent, otherwise there exists $T_1v = 0$ where v is linear combination of $v_1, \ldots, v_m$. This means $v \in \text{null } T_1 = \text{null } T_2$ so $T_2v = 0$, which leads to $T_2v_1, \ldots, T_2v_m$ being linearly dependent, a contradiction. Thus, $T_1v_1, \ldots, T_1v_m$ is linearly independent so it can be extended to basis $T_1v_1, \ldots, T_1v_m, u_1, \ldots u_n$ of W.

Let $S \in \mathcal{L}(W)$ be linear map so $S(T_2v_i) = T_1v_i$ for $1 \le i \le m$ and $Sw_j = u_j$ for $1 \le j \le n$. This follows S is surjective so S is invertible.

Now we prove $ST_2 = T_1$. Indeed, for any $v \in V$ then $T_2v = \sum_{i=1}^m \alpha_i T_2v_i$ so $v - \sum_{i=1}^m \alpha_i v_i \in$ null T_2 = null T_1 so

$$T_1v = T_1\left(\sum_{i=1}^m \alpha_i v_i\right) = \sum_{i=1}^m \alpha_i T_1v_i = \sum_{i=1}^m \alpha_i ST_2v_i = ST_2\left(\sum_{i=1}^m \alpha_i V_I\right) = S(T_2v).$$

exer:3D:5

5. If there exists such invertible operator S then for any $u \in \text{range } T_1, u = T_1v = T_2(Sv)$ so $u \in \text{range } T_2$. Similarly, if $u \in \text{range } T_2$ then $u \in \text{range } T_1$. Thus, range T_1 = range T_2.

   Conversely, for $v_1, \ldots, v_m$ as basis of V, since range T_1 = range T_2 there exists u_i so $T_1v_i = T_2u_i$ for all $1 \le i \le m$. According to theorem 5.1.1, there exists a linear map $S \in \mathcal{L}(V, V)$ so $Sv_i = u_i$ for all $1 \le i \le m$. This follows $T_2Sv_i = T_2u_i = T_1v_i$ for all $1 \le i \le m$, which leads to $T_2Sv = T_1v$ for all $v \in V$. On the other hand, since $v_1, \ldots, v_m$ is linearly independent so $u_1, \ldots, u_m$ is linearly independent so it is basis of V. Hence, S is surjective so S is invertible.

exer:3D:6

6. If there exists such R, S. Let $v_1, \ldots, v_m$ be basis of null T_1 then $ST_2Rv_i = 0$. Since S is invertible so S is injective, which means $T_2Rv_i = 0$. Thus, $Rv_i \in \text{null } T_2$ and $Rv_1, \ldots, Rv_m$ is linearly independent so dim null $T_2 \ge$ dim null T_1. Since R is invertible so R is surjective so let $Ru_1, \ldots, Ru_n$ be basis of null T_2. Hence $S(T_2Ru_i) = S(0) = 0$ so $T_1u_i = 0$ so $u_i \in \text{null } T_1$. Since $u_1, \ldots, u_n$ is linearly independent so dim null $T_2 \le$ dim null T_1. Thus, dim null T_1 = dim null T_2.

   Conversely, if dim null T_1 = dim null T_2: Let $v_1, \ldots, v_m$ be basis of V where $v_{n+1}, \ldots, v_m$ ($n \le m$) is basis of null T_1. Let $u_1, \ldots, u_m$ be another basis of V where $u_{n+1}, \ldots, u_m$ is basis of null T_2. This folows that two lists $T_1v_1, \ldots, T_1, v_n$ and $T_2u_1, \ldots, T_2u_n$ are linearly independent so they can be extended to two bases of W, which are $T_1v_1, \ldots, T_1v_n, w_1, \ldots, w_k$ and $T_2u_1, \ldots, T_2u_n, z_1, \ldots, z_k$.

   According to theorem 5.1.1, there exists a linear map $R \in \mathcal{L}(V)$ so $Rv_i = u_i$ for $1 \le i \le m$. Since $u_1, \ldots, u_m$ is basis of V so R is surjective so R is invertible. There also exists linear map $S \in \mathcal{L}(W)$ so $S(T_2u_i) = T_1v_i$ for $1 \le i \le n$ and $Sz_j = w_j$ for $1 \le j \le k$. We can see that S is also surjective so S is invertible.

Now we prove $T_1 = ST_2R$. Indeed, for any $v = \sum_{i=1}^m \alpha_i v_i$ in V, we have

$$\begin{aligned} ST_2Rv &= ST_2\left(\sum_{i=1}^m \alpha_i u_i\right) = ST_2\left(\sum_{i=1}^n \alpha_i u_i\right) \text{ (since } T_2u_j = 0 \text{ for } m \geq j > n), \\ &= S\left(\sum_{i=1}^n \alpha_i T_2 u_i\right) = \sum_{i=1}^n \alpha_i T_1 v_i, \\ &= \sum_{i=1}^m \alpha_i T_1 v_i = T_1 v \text{ (since } T_1 v_j = 0 \text{ for } m \geq j > n). \end{aligned}$$

exer:3D:7

7. (a) For any $T_1, T_2 \in E$ then we have $(T_1 + T_2)(v) = T_1v + T_2v = 0$ so $T_1 + T_2 \in E$ and $(\lambda T_1)(v) = \lambda(T_1 v) = 0$ so $\lambda T_1 \in E$. We also have $0 \in E$ since $0(v) = 0$. Thus, E is subspace of V.

   (b) Let $v, v_1, \dots, v_n$ be basis of V and $w_1, \dots, w_m$ be basis of W. We prove that $\mathcal{M}$ with respect to these bases is an isomorphism between two vector spaces E and $D = \{A \in \mathbf{F}^{m,n} : A_{\cdot,1} = 0\}$ (need to prove D is vector space first, which is not hard).

   Indeed, $\mathcal{M}$ is linear. If $\mathcal{M}(T) = 0$ then $T = 0$. Thus, $\mathcal{M}$ is injective. For any $A \in D$, let T be the linear map from E to D so $Tv_k = \sum_{i=1}^m A_{i,k} w_i$ for $1 \leq k \leq n$ and $Tv = 0$. This follows $T \in E$. Thus, $\mathcal{M}$ is surjective, which follows $\mathcal{M}$ is invertible.

   Hence, according to theorem 5.6.2, dim E = dim D. One can easily verify that dim D = mn = (dim $V - 1$)dim W so dim E = (dim $V - 1$)dim W.

exer:3D:8

8. Since T is surjective so range $T = W$ so W is finite-dimensional. Let $v_1, \dots, v_m$ be basis of V where $v_{n+1}, \dots, m$ $(n \leq m)$ is basis of null T. Let $U = \text{span}\,(v_1, \dots, v_n)$. For any $w \in W$, there exists $v = \sum_{i=1}^m \alpha_i v_i$ so $Tv = w$ so $u = v - \sum_{i=n+1}^m \alpha_i v_i$ is in U and $Tu = Tv = w$. Thus, $T|_U$ is surjective. If $Tu = Tv$ for $u, v \in U$ then $u - v \in \text{null}\, T = \text{span}\,(v_{n+1}, \dots, v_m)$ so $u - v = 0$ since $v_1, \dots, v_m$ is linearly independent. Thus, $T|_U$ is injective. Thus, $T|_U$ is an isomorphism of U onto W.

exer:3D:9

9. If both S, T are invertible then $T^{-1}S^{-1}$ is inverse of ST so ST is invertible. Conversely, if ST is invertible then there is an inverse X of ST. For any $v \in V$, there exists $TXv \in V$ so $S(TXv) = v$. Hence, S is surjective so according to theorem 5.6.4, S is invertible. On the other hand, if $Tu = Tv$ for $u, v \in V$ then $XS(Tu) = XS(Tv)$ but $u = XSTu, v = XSTv$ so $u = v$. Thus, T is injective so according to theorem 5.6.4, T is invertible.

exer:3D:10

10. If $ST = I$ then for any $v \in V$, we have $S(Tv) = v$ so S is surjective so according to theorem 5.6.4, S is invertible. Let S^{-1} be inverse of S then we have $T = IT = (S^{-1}S)T = S^{-1}(ST) = S^{-1}$. Thus, T is inverse of S so $TS = I$. Similarly, if $TS = I$ then $ST = I$.

exer:3D:11

11. According to exercise 10 (3D) if $S(TU) = I$ then $(TU)S = T(US) = I$ then $(US)T = I$ so $T^{-1} = US$.

exer:3D:12

12. Define $T, U, S \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ so $Uv = vx^3$ for any $v \in \mathcal{P}(\mathbf{R})$; $Tx^i = x^{i-2}$ for $i \geq 2$ and $Tx = x, T1 = 1$; $Sx^i = x^{i-1}$ for $i \geq 1$ and $S1 = 1$. With this, we have $STU = I$. However, if T is not invertible since T is not injective as $T(2x^2 + x^3) = T(2 + x^3) = 2 + x$.

exer:3D:13 13. Note that $RST \in \mathcal{L}(V)$ and RST is surjective so according to theorem 5.6.4, RST is invertible operator. According to exercise 9 (3D), since RST is invertible so R and ST are also invertible, which follows S, T, R are invertible. Thus, S is injective.

exer:3D:14 14. T is linear. We have $Tu = Tv$ then $\mathcal{W}(u) = \mathcal{W}(v)$ so $u = v$. Thus, T is injective. For any $A \in \mathbf{F}^{n,1}$ there exists a $v \in V$ so $Tv = A$. Thus, $\mathcal{M}$ is isomorphism of V onto $\mathbf{F}^{n,1}$.

exer:3D:15 15. Let $A = \mathcal{M}(T)$ with respect to standard bases of both $\mathbf{F}^{n,1}, \mathbf{F}^{m,1}$. Then for any $v \in \mathbf{F}^{n,1}$, we have $\mathcal{M}(v) = v$ and $Tv = \mathcal{M}(Tv)$. Hence, according to theorem 5.6.3, we have $Tv = \mathcal{M}(Tv) = \mathcal{M}(T)\mathcal{M}(v) = Av$.

exer:3D:16 16. Let $v_1, \ldots, v_n$ be basis of V. First, we prove that for any $1 \leq i \leq n$, there exists $C_i \in \mathbf{F}$ so $Tv_i = C_i v_i$. Indeed, we can choose linear map S so $Sv_i = v_i$ and $Sv_j = 2v_j$ for $1 \leq j \leq n, j \neq i$. This follows $STv_i = T(Sv_i) = Tv_i$. Hence, if $Tv_i = \sum_{k=1}^n \alpha_k v_k$ then

$$\sum_{k=1}^n \alpha_k v_k = Tv_i = STv_i = S\left(\sum_{k=1}^n \alpha_k v_k\right) = \alpha_i v_i + 2\sum_{k \neq i} \alpha_k v_k.$$

This follows $\sum_{k \neq i} \alpha_k v_k = 0$, which happens when $\alpha_k = 0$ for $k \neq i$. Thus, $Tv_i = \alpha_i v_i = C_i v_i$.

Next, we prove that for any two $1 \leq j, i \leq n$ so $j \neq i$ then $C_i = C_j = C$. Indeed, pick an invertible operator S so $Sv_i = v_j$ then $TSv_i = Tv_j = C_j v_j$ but $STv_i = S(C_i v_i) = C_i v_j$. Hence, $C_i v_j = C_j v_j$ so $C_i = C_j = C$. Thus, $Tv_i = Cv_i$ for all $1 \leq i \leq n$, or T is scalar multiple of I.

exer:3D:17 17. (For this, it's better to think linear maps as matrices) Let $v_1, \ldots, v_n$ be basis of V. Denote linear map $T_{i,j}$ as $Tv_j = v_i, Tv_k = 0$ for all $k \neq j, 1 \leq k \leq n$ (image this as n-by-n matrix with 0 in all entries except for 1 in (i,j)-th entry).

If all $T \in \mathcal{E}$ is 0 then $\mathcal{E} = \{0\}$. If there exists $T \neq 0$, WLOG $Tv_1 = \sum_{i=1}^n \alpha_i v_i$ with $\alpha_1 \neq 0$. We find $T_{i,1}(TT_{1,1}) \in \mathcal{E}$. Note that $T_{i,1}TT_{1,1} = \alpha_1 T_{i,1}$ so $T_{i,1} \in \mathcal{E}$ for all $1 \leq i \leq n$. Since $T_{i,1} \in \mathcal{E}$ so $T_{i,j} = T_{i,1}T_{1,j} \in \mathcal{E}$ for all $1 \leq j \leq n$. Thus, $T_{i,j} \in \mathcal{E}$ for all $1 \leq i, j \leq n$. Since $T_{1,1}, \ldots, T_{n,n}$ is linearly independent and $\dim V = n^2$ so $\mathcal{E} = V$.

exer:3D:18 18. Define a map T from V onto $\mathcal{L}(\mathbf{F}, V)$ so for any $v \in V$, $Tv = R_v \in \mathcal{L}(\mathbf{F}, V)$ so $R_v 1 = v$. it's not hard to verify that T is linear. If $Tu = Tv$ then $R_v = R_u$ so $R_v(1) = R_u(1)$ so $u = v$. Thus, T is injective. For any $R \in \mathcal{L}(\mathbf{F}, V)$, then there is $R(1) \in V$ so $V(R(1)) = R$. Thus, T is surjective. In conclusion, T is an isomorphism so V and $\mathcal{L}(\mathbf{F}, V)$ are isomorphic vector spaces.

exer:3D:19 19. (a) For any $q \in \mathcal{P}(\mathbf{R})$, which we can restrict to $q \in \mathcal{P}_m(\mathbf{R})$. Define $T' : \mathcal{P}_m(\mathbf{R}) \to \mathcal{P}_m(\mathbf{R})$ so $T'p = Tp$ for $p \in \mathcal{P}_m(\mathbf{R})$. Since $\deg Tp \leq \deg p = m$ so T' is indeed an operator on finite-dimensional $\mathcal{P}_m(\mathbf{R})$. Since T is injective so T' is also injective, which follows T' is invertible. Hence, for $q \in \mathcal{P}_m(\mathbf{R})$ there exists $p \in \mathcal{P}_m(\mathbf{R})$ so $T'p = Tp = q$. This follows T is surjective.

(b) Assume the contrary, there is non-zero p so $\deg Tp < \deg p$. Let $Tp = q$. From (a), there exists $r \in \mathcal{P}(\mathbf{R})$ so $\deg r \leq \deg q$ and $Tr = q$. This follows $Tp = Tr$ so $p = r$. However $\deg p > \deg Tp = \deg q \geq \deg r$, a contradiction. Thus, we must have $\deg Tp = \deg p$ for any non-zero $p \in \mathcal{P}(\mathbf{R})$.

exer:3D:20 20. Denote n-by-n matrix A with $A_{i,j}$ is the (i,j)-th entry. Define a linear map $T \in \mathcal{L}(\mathbf{F}^{n,1})$ so $Tv = Av$. Hence, (a) is equivalent to T being injective and (b) is equivalent to T being surjective. Thus, (a) is equivalent to (b) according to theorem 5.6.4. theo:3.69:3D

5.8. 3E: Products and Quotients of Vector Spaces

theo:3.85:3E **Theorem 5.8.1** (3.85) Suppose U is a subspace of V and $u, v \in V$. Then the following are equivalent:

(a) $v - w \in U$.

(b) $v + U = w + U$.

(c) $(v + U) \cap (w + U) \neq 0$.

theo:3.89:3E **Theorem 5.8.2** (3.89) Suppose V is finite-dimensional and U is subspace of V. Then $\dim V/U = \dim V - \dim U$.

theo:3.91:3E **Theorem 5.8.3** (3.91) Suppose $T \in \mathcal{L}(V, W)$. Define $\tilde{T} : V/(\text{null } T) \to W$ by $\tilde{T}(v + \text{null } T) = Tv$. Then:

(a) $\tilde{T}$ is a linear map from $V/(\text{null } T)$ to W.

(b) $\tilde{T}$ is injective,

(c) $\text{range } \tilde{T} = \text{range } T$,

(d) $V/(\text{null } T)$ is isomorphic to range T.

Remark 5.8.4. Be careful with vector space V/U, because there are more than one $v \in V$ to give the same affine subset $v + U$ of V. So if we define a map T from V/U to some W, be sure that T makes senses, i.e. for $v + U = u + U$ then $T(v + U) = T(u + U)$.

5.9. Exercises 3E

exer:3E:1 1. If T is a linear map then for any two $u, v \in V$ we have $(u, Tu) + (v, Tv) = (u + v, T(u + v)) \in$ graph of T and $\lambda(u, Tu) = (\lambda u, T\lambda u) \in$ graph of T. We also have $(0, T0) \in$ graph of T so graph of T is a subspace of $V \times W$.

Conversely, if graph of T is a subspace of $V \times W$. We have $(u, Tu) + (v, Tv) = (u + v, Tu + Tv) \in$ graph of T so $Tu + Tv = T(u + v)$ (otherwise if $Tu + Tv \neq T(u + v)$ then function

T maps $u+v$ to two different results, a contradiction). Similarly, $\lambda(u, Tu) = (\lambda u, \lambda Tu) \in$ range T so $\lambda Tu = T(\lambda u)$. Thus, T is a linear map.

exer:3E:2 2. Let $(v_{1,1}, v_{1,2} \ldots, v_{1,m}), (v_{2,1}, \ldots, v_{2,m}), \ldots, (v_{k,1}, \ldots, v_{k,m})$ be basis of $V_1 \times V_2 \times \cdots \times V_m$. Assume the contrary that V_1 is infinite-dimensional. then there exists $v \in V_1$ so $v \notin$ span $(v_{1,1}, v_{2,1}, \ldots, v_{k,1})$. This follows $(v, u_2, \ldots, u_m)$ for any $u_i \in V_i$ is not a linear combination of $(v_{i,1}, \ldots, v_{i,m})$, $1 \le i \le k$, a contradiction. Thus, V_i is finite-dimensional.

exer:3E:3 3. Let $V = \mathcal{P}(\mathbf{R}), U_1 = \text{span }(1, x, x^2)$ and $U_2 = \{p \in \mathcal{P}(\mathbf{R}) : \deg p \ge 2\}$. We find $U_1 + U_2$ is not a direct sum since $1 + x + 2x^2 + x^3 = (1 + x + x^2) + (x^2 + x^3) = (1 + x) + (2x^2 + x^3)$. Define a linear map T from $U_1 \times U_2$ onto $U_1 + U_2$ so $T(u_1, u_2) = u_1 + xu_2$. We can find that T is surjective. If $T(u_1, u_2) = T(v_1, v_2)$ then $u_1 + xu_2 = v_1 + xv_2$. Note that $\deg u_1 < \deg xu_2$ for any $u_1 \in U_1, u_2 \in U_2$ so this only happens when $u_1 = v_1, u_2 = v_2$. Thus, T is injective which follows T is an isomorphism from $U_1 \times U_2$ onto $U_1 + U_2$.

exer:3E:4 4. Define a map S from $\mathcal{L}(V_1 \times \cdots \times V_m, W)$ onto $\mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ so for any $T \in \mathcal{L}(V_1 \times \cdots \times V_m, W)$, $S(T) = (R_1, \ldots, R_m)$ where R_i are map from V_i onto W so: for any $v = (0, \ldots, 0, v_i, 0, \ldots, 0) \in V_1 \times \cdots \times V_m$ then $R_i v_i = Tv = w \in W$. It's not hard to check that R_i is a linear map.

Now we prove that S is a linear map. Say $S(T_1) = (R_1, \ldots, R_m)$ and $S(T_2) = (U_1, \ldots, U_m)$ then $S(T_1) + S(T_2) = (R_1 + U_1, \ldots, R_m + U_m)$. On the other hand, for any $1 \le i \le m$, we have for any $v_i \in V_i$ then $(R_i + U_i)v_i = R_i v_i + U_i v_i = T_1 v + T_2 v = (T_1 + T_2)v$ where $v = (0, \ldots, 0, v_i, 0, \ldots, 0)$. This follows $S(T_1 + T_2) = (R_1 + U_1, \ldots, R_m + U_m) = S(T_1) + S(T_2)$.

Similarly, for any $(R_1, \ldots, R_m) \in \mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ we can define respective $T \in \mathcal{L}(V_1 \times \cdots \times V_m, W)$. Thus, S is surjective. If $S(T_1) = S(T_2)$ then $(R_1, \ldots, R_m) = (U_1, \ldots, U_m)$ so $R_i = U_i$ or $T_1 v_i = T_2 v_i$ where $v_i = (0, \ldots, 0, v_i, 0, \ldots, 0)$ for all $1 \le i \le m$. This leads to $T_1 = T_2$. Thus, S is injective so S is an isomorphism from $\mathcal{L}(V_1 \times \cdots \times V_m, W)$ onto $\mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ so these two vector spaces are isomorphic.

exer:3E:5 5. Similarly to previous exercise 4, define a linear map S from $\mathcal{L}(V, W_1 \times \cdots \times W_m)$ onto $\mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$ so for $T \in \mathcal{L}(V, W_1 \times \cdots \times W_m)$, $S(T) = (R_1, \ldots, R_m)$ where $R_i \in \mathcal{L}(V, W_i)$ so: if $Tv = (w_1, \ldots, w_m)$ then $R_i v_i = w_i$.

exer:3E:6 6. Define a linear map T from V^n onto $\mathcal{L}(\mathbf{F}^n, V)$ so $T(v_1, \ldots, v_n) = S \in \mathcal{L}(\mathbf{F}^n, V)$ so $S(0, \ldots, 0, 1_i, 0, \ldots, 0) = v_i$. T is obviously surjective. If $T(v_1, \ldots, v_n) = T(u_1, \ldots, u_n)$ then $S_v = S_u$ so $S_v(0, \ldots, 0, 1_i, 0, \ldots, 0) = S_u(0, \ldots, 0, 1_i, 0, \ldots, 0)$ which means $v_i = u_i$. This follows $(v_1, \ldots, v_n) = (u_1, \ldots, u_n)$. Thus, T is injective so two vector spaces V^n and $\mathcal{L}(\mathbf{F}^n, V)$ are isomorphic.

exer:3E:7 7. For any $u \in U$, we have $v + u \in v + U = x + W$ so $v + u - x \in W$ which according to theorem 5.8.1 follows $v + u + W = x + W = v + U$ so $u + W = U$. Since $u \in U$ so $W \subseteq U$.

Similarly, for any $w \in W$ then $w + x \in x + W = v + U$ so $w + x - v \in U$ so $w + x + U = v + U = x + W$ so $w + U = W$. Thus, $U \subseteq W$. In conclusion, we find $U = W$.

exer:3E:8 8. If A is an affine subset of V then A is $r+U$ for some $r \in V$ and some subspace U of V. For $v, w \in A$ then there exists $u_1, u_2 \in U$ so $r+u_1 = v, r+u_2 = w$. We find $\lambda v + (1-\lambda)w = r + \lambda u_1 + (1-\lambda)u_2 \in A$ for any $\lambda \in \mathbf{F}$.

Conversely, if $\lambda v + (1-\lambda)w \in A$ for all $v, w \in A$ and all $\lambda \in \mathbf{F}$: Pick arbitrary $w \in A$, we prove that $U = \{u - w : u \in A\}$ is a subspace of V. Indeed, for any $x_1, x_2 \in U$ then $x_1 = u_1 - w, x_2 = u_2 - w$ for $u_1, u_2 \in A$. Pick $\lambda = 2$ then $2u_1 - w \in A, 2u_2 - w \in A$. Pick $\lambda = 1/2$ then $1/2(2u_1 - w) + 1/2(2u_2 - w) \in A$ or $u_1 + u_1 - w \in A$. This follows $u_1 + u_2 - 2w \in u$ or $x_1 + x_2 \in U$ for any $x_1, x_2 \in U$. On the other hand, $\lambda(2u_1 - w) + (1-\lambda)w \in A$ or $2\lambda u_1 \in A$ for any $u_1 \in A$. Thus, U is a subspace of V. This follows $A = w + U$ so A is an affine subset of V.

exer:3E:9 9. If $A_1 \cap A_2$ is nonempty then for any $x, y \in A_1 \cap A_2$, according to exercise 8, $\lambda x + (1-\lambda)y \in A_1$ and $\lambda x + (1-\lambda)y \in A_2$ so $\lambda x + (1-\lambda)y \in A_1 \cap A_2$ for any $\lambda \in \mathbf{F}, x, y \in A_1 \cap A_2$. Thus, according to exercise 8, we find $A_1 \cap A_2$ is an affine subset of V.

exer:3E:10 10. This is deduced from exercise 9.

exer:3E:11 11. (a) For any $x, y \in A$ where $x = \sum_{i=1}^m \lambda_i v_i, y = \sum_{i=1}^m \beta_i v_i$ then for any $\alpha \in \mathbf{F}$, we have $(1-\alpha)x + \alpha y = \sum_{i=1}^m v_i \left[\lambda_i(1-\alpha) + \alpha\beta_i\right]$. Since

$$\sum_{i=1}^m \lambda_i(1-\alpha) + \alpha\beta_i = (1-\alpha)\sum_{i=1}^m \lambda_i + \alpha\sum_{i=1}^m \beta_i = 1$$

so according to exercise 8, A is an affine subset of V.

(b) Let U be affine subset of V that contains $v_1, \ldots, v_m$. Hence, $\lambda_1 + \lambda_2 v_2 \in U$ for any $\lambda_1, \lambda_2 \in \mathbf{F}; \lambda_1 + \lambda_2 = 1$. This follows $\lambda_3 v_3 + (1-\lambda_3)(\lambda_1 v_1 + \lambda_2 v_2) \in U$ for any $\lambda_3 \in \mathbf{F}$. Since $\lambda_3 + (1-\lambda_3)\lambda_1 + (1-\lambda_3)\lambda_2 = 1$ so we can say that $\lambda_3 v_3 + \lambda_2 v_2 + \lambda_1 v_1 \in U$ for any $\lambda_1, \lambda_2, \lambda_3 \in \mathbf{F}$ so $\lambda_1 + \lambda_2 + \lambda_3 = 1$. Similarly, $\lambda_4 v_4 + (1-\lambda_4)(\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3) \in U$. With this, we will eventually obtain $\sum_{i=1}^m \lambda_i v_i \in U$ for $\sum_{i=1}^m \lambda_i = 1$ or U contains A.

(c) We find $v + 0 \in A$ so $v = \sum_{i=1}^m \alpha_i v_i$ with $\sum_{i=1}^m \alpha_i = 1$. This follows for any $u \in U$ then $u = \sum_{i=1}^m \beta_i v_i$ with $\sum_{i=1}^m \beta_i = 0$ or $u = \sum_{i=1}^{m-1} \beta_i(v_i - v_m)$ since $\beta_m = -\beta_1 - \ldots - \beta_{m-1}$. Thus, U is a subspace of span $(v_1 - v_m, v_2 - v_m, \ldots, v_{m-1} - v_m)$ so $\dim U \le m-1$.

exer:3E:12 12. Let $v_1 + U, v_2 + U, \ldots, v_m + U$ be basis of V/U then $v_1, \ldots, v_m$ is linearly independent and $\sum_{i=1}^m \alpha_i v_i \notin U$ for any not-all-zero $\alpha_i \in \mathbf{F}$.

Define a linear map T from V onto $U \times (V/U)$ so $Tv = (v - \sum_{i=1}^m \alpha_i v_i, v + U)$ where $v + U = (\sum_{i=1}^m \alpha_i v_j) + U$. According to theorem 5.8.1 then $v - \sum_{i=1}^m \alpha_i v_i \in U$ so $Tv \in U \times (V/U)$.

For any $(w, v+U) \in U \times (V/U)$ there exists $\alpha_1, \ldots, \alpha_m \in \mathbf{F}$ so $v + U = (\sum_{i=1}^m \alpha_i v_i) + U$. By letting $u = w + \sum_{i=1}^m \alpha_i v_i \in V$ then we have $Tu = (w, v+U)$. Thus, T is surjective.

If $Tv_1 = Tv_2$ or $(w_1, x_1 + U) = (w_2, x_2 + U)$ then $w_1 = w_2$ and $x_1 + U = x_2 + U = (\sum_{i=1}^m \alpha_i v_i) + U$. Hence,

$$v_1 = w_1 + \sum_{i=1}^m \alpha_i v_i = w_2 + \sum_{i=1}^m \alpha_i v_i = v_2.$$

Thus, T is injective. In conclusion, vector spaces V is isomorphic to $U \times (V/U)$.

exer:3E:13 13. Since $v_1+U, v_2+U, \ldots, v_m+U$ is basis of V/U so if $\sum_{i=1}^m \alpha_i(v_i+U) = 0+U$ then $\alpha_i = 0$ for all $1 \leq i \leq m$, i.e. if $\sum_{i=1}^m \alpha_i v_i \neq 0$ then $\sum_{i=1}^m \alpha_i v_i \notin U$. We also find that $v_1, \ldots, v_m$ is linearly independent. From all the above, we deduce $v_1, \ldots, v_m, u_1, \ldots, u_n$ is a linearly dependent list. According to theorem 5.8.2 then $m+n = \dim V/U + \dim U = \dim V$ so linearly independent list $v_1, \ldots, v_m, u_1, \ldots, u_n$ is basis of V.

exer:3E:14 14. (a) If $(x_1, x_2, \ldots) \in U$ then there exists $M_1 \in \mathbf{N}$ so for all $i \geq M_1$ then $x_i = 0$. Similarly, if $(y_1, y_2, \ldots) \in U$ then there exists M_2. Thus, for $(x_1+y_1, x_2+y_2, \ldots)$, for all $i \geq \max\{M_1, M_2\}$ then $x_i + y_i = 0$, which means $(x_1+y_1, x_2+y_2, \ldots) \in U$. Similarly $\lambda(x_1, \ldots) \in U$. And $(0, 0, \ldots) \in U$ so U is subspace of $\mathbf{F}^\infty$.

(b) Denote p_i as i-th prime number and $v_i \in \mathbf{F}^\infty$ so for any $k \in \mathbf{N}, k \geq 1$ then the kp_i-th coordinate of v_i is 2, the rest of the coordinates of v_i is 1. We will prove that prove that $v_1+U, \ldots, v_m+U$ is linearly independent for any $m \geq 1$, which can follow that $\mathbf{F}^\infty/U$ is infinite-dimensional according to exercise 14 (2A).

Indeed, consider $(\sum_{i=1}^m \alpha_i v_i) + U$ is equal to U when $v = \sum_{i=1}^m \alpha_i v_i \in U$ according to theorem 5.8.1, which means there exists $M \in \mathbf{N}$ so for all $i \geq M$, i-th coordinate of v is 0. According to notation of v_i, the x-th coordinate of v where

- $x \equiv 1 \pmod{p_1 \cdots p_m}$ is $\sum_{i=1}^m \alpha_i$.
- $x \equiv 0 \pmod{p_j}, x \equiv 1 \pmod{\prod_{i \neq j} p_i}$ is $2\alpha_j + \sum_{i \neq j} \alpha_i$.

If we pick $x \geq M$ then from the above, we find that $\sum_{i=1}^m \alpha_i = 2\alpha_j + \sum_{i \neq j} \alpha_i = 0$ for every $1 \leq j \leq m$. Hence, we find $\alpha_i = 0$ for all $1 \leq i \leq m$. In other words, $v_1+U, v_2+U, \ldots, v_m+U$ is linearly independent in $\mathbf{F}^\infty/U$, as desired.

exer:3E:15 15. According to theorem 5.8.3, $V/(\text{null } \varphi)$ is isomorphic to range φ. We find that range $\varphi = \mathbf{F}$ since $\varphi \neq 0$ so for any $c \in \mathbf{F}$, there exists $\lambda \in \mathbf{F}$ so $\lambda\varphi(1) = c$. Hence, $\dim \text{range } \varphi = \dim \mathbf{F} = 1$, which follows $\dim V/(\text{null } \varphi) = 1$.

exer:3E:16 16. Since $\dim V/U = 1$ so there exists $u \in V, u \notin U$. Consider $v \neq u, v \notin U, v \in V$ then $v + U = \lambda u + U$ for some $\lambda \in \mathbf{F}$, hence according to theorem 5.8.1, $v = \lambda u + x$ for some $x \in U$.

Define a linear map $\varphi \in \mathcal{L}(V, \mathbf{F})$ so $\varphi(u) = 1$, $\varphi(v) = 0$ for all $v \notin U$, $\varphi(v) = \lambda\varphi(u)$ if $v + U = \lambda u + U$. Consider $\varphi v = 0$ for some $v \in V$. If $v \notin U$ then $v = \lambda u + x$ for some $\lambda \in \mathbf{F}, x \in U$ so $\varphi(v) \neq 0$, a contradiction. Thus, only $v \in U$ will give $\varphi v = 0$, which means null $\varphi = U$.

exer:3E:17 17. Let $v_1+U, v_2+U, \ldots, v_m+U$ be basis of U. For any $v \in V$ then there exists uniquely $\alpha_1, \ldots, \alpha_m \in \mathbf{F}$ so $v + U = (\sum_{i=1}^m \alpha_i v_i) + U$. According to theorem 5.8.1, $v - \sum_{I=1}^m \alpha_i v_i = u \in U$. Hence, every $v \in V$ can be represented uniquely as $u+x$ where $x \in \text{span }(v_1, \ldots, v_m)$. Thus, $V = U \oplus \text{span }(v_1, \ldots, v_m)$.

exer:3E:18 18. If there exists such linear map S then $0 = T(0) = S(\pi 0) = S(U)$. For any $u \in U$ then $T(u) = S(\pi u) = S(U) = 0$ so $u \in \text{null } T$. Thus, $U \subset \text{null } T$.

Conversely, if $U \subset \text{null } T$. Define $S : V/U \to W$ so $S(v+U) = T(v)$. We show that the definition of S makes sense, suppose $u, v \in V$ so $u + U = v + U$ then according to theorem 5.8.1, $v - u \in U$ so $v - u \in \text{null } T$ so $Tv = Tu$.

exer:3E:19 19. Disjoint union is analogous to direct sum: If number of elements of $A_1 \cup A_2 \cup \ldots \cup A_m$ is equal to sum of number of elements of A_i, then $A_1 \cup \ldots \cup A_m$ is a disjoint union.

exer:3E:20 20. (a) We have $\Gamma(S_1 + S_2) = (S_1 + S_2) \circ \pi = S_1 \circ \pi + S_2 \circ \pi = \Gamma(S_1) + \Gamma(S_2)$. Similarly, $\Gamma(\lambda S_1) = (\lambda S_1) \circ \pi = \lambda \Gamma(S_1)$. Thus, Γ is a linear map.

(b) If $\Gamma(S_1) = \Gamma(S_2)$ then $S_1 \circ \pi = S_2 \circ \pi$ or $S_1(v+U) = S_2 = (v+U)$ or $(S_1 - S_2)(v+U)$ for any $v + U \in V/U$. Hence, $S_1 = S_2$ so Γ is injective.

(c) For any $T \in \{T \in \mathcal{L}(V, W) : Tu = 0 \text{ for every } u \in U\}$, according to exercise 18, there exists $S \in \mathcal{L}(V/U, W)$ so $\Gamma(S) = T = S \circ \pi$. Thus, range $\Gamma = \{T \in \mathcal{L}(V, W) : Tu = 0 \text{ for every } u \in U\}$.

5.10. 3F: Duality

theo:3.95:3F **Theorem 5.10.1** (3.95) Suppose V is finite-dimensional. Then $V' = \mathcal{L}(V, \mathbf{F})$ is also finite-dimensional and $\dim V' = \dim V$.

theo:3.105:3F **Theorem 5.10.2** (3.105) Suppose $U \subset V$ then U^0 is a subspace of V'.

theo:3.106:3F **Theorem 5.10.3** (3.106) Suppose V is finite-dimensional and U is subspace of V. Then $\dim U + \dim U^0 = \dim V$.

theo:3.107:3F **Theorem 5.10.4** (3.107) Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then:

(a) $\text{null } T' = (\text{range } T)^0$ (don't need V, W being finite-dimensional).

(b) $\dim \text{null } T' = \dim \text{null } T + \dim W - \dim V$.

theo:3.108,3.110:3F **Theorem 5.10.5** (3.108,3.110) Suppose V, W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then T is injective if and only if T' is surjective and T is surjective if and only if T' is injective.

theo:3.109:3F **Theorem 5.10.6** (3.109) Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then:

(a) $\dim \text{range } T' = \dim \text{range } T$;

(b) $\text{range } T' = \text{null } T$.

theo:3.117:3F **Theorem 5.10.7** (3.117) Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then dim range T equals to column rank of $\mathcal{M}(T)$.

5.11. Exercises 3F

exer:3F:1 1. Let $\varphi \in \mathcal{L}(V, \mathbf{F})$. If φ is not a zero map then there exists $v \in V$ so $\varphi v = c \neq 0$. Hence, for any $\lambda \in \mathbf{F}$ then $\varphi(\lambda/c \cdot v) = \lambda$ so φ is surjective.

exer:3F:2 2. Linear functionals on $\mathbf{R}^{[0,1]}$: $\varphi_1(f) = f(0.5), \varphi_2(f) = f(0.1), \varphi_3(f) = f(0.2)$.

exer:3F:3 3. Pick $\varphi \in V'$ so $\varphi(v) = 1$, value of $\varphi(u)$ for $u \notin \{av : a \in \mathbf{F}\}$ can be picked arbitrarily.

exer:3F:4 4. Extend basis $v_1, \ldots, v_m$ of U to basis $v_1, \ldots, v_n$ $(m < n)$ of V. Define φ from V onto $\mathbf{F}$ so $\varphi(v_i) = 0$ for $1 \leq i \leq m$, and $\varphi(v_j) = 1$ for $m < j \leq n$. Hence, $\varphi(u) = 0$ for every $u \in U$.

exer:3F:5 5. It is a consequence of exercise 4 (3E) when $W = \mathbf{F}$.

exer:3F:6 6. (a) If $v_1, \ldots, v_m$ spans V: Consider $\varphi \in V'$ so $\Gamma(\varphi) = 0$ then $\varphi(v_i) = 0$ for all $1 \leq i \leq m$, which leads to $\varphi(v) = 0$ for all $v \in V$ since $v_1, \ldots, v_m$ spans V. Hence, $\varphi = 0$, i.e. Γ is injective.

   Conversely, if Γ is injective. If $v_1, \ldots, v_m$ does not span V then there exists $v_{m+1} \notin$ span $(v_1, \ldots, v_m), v_{m+1} \in V$. Hence, there exists $\varphi \in V'$ so $\varphi(v_{m+1}) = 1$ and $\varphi(v) = 0$ for all $v \in V, v \notin \{av_{m+1} : a \in \mathbf{F}\}$. We find $\Gamma(\varphi) = 0$ but $\varphi \neq 0$, a contradiction to injectivity of Γ. Thus, $v_1, \ldots, v_m$ spans V.

   (b) If $v_1, \ldots, v_m$ is linearly independent then for any $(x_1, \ldots, x_m) \in \mathbf{F}^m$, we can construct $\varphi \in V'$ so $\varphi(v_i) = x_i$. Hence, Γ is surjective.

   If Γ is surjective but $v_1, \ldots, v_m$ is linearly dependent, i.e. WLOG $v_m = \sum_{i=1}^{m-1} \alpha_i v_i$. Hence, $\Gamma(\varphi) = \left(\varphi(v_1), \ldots, \varphi(v_{m-1}), \sum_{i=1}^{m-1} \alpha_i \varphi_i(v_i)\right)$, which does not cover all $\mathbf{F}^m$. Thus, $v_1, \ldots, v_m$ is linearly independent.

exer:3F:7 7. Because $\varphi_j(x^j) = \frac{(x^j)^{(}j)(0)}{j!} = 1$ and $\varphi_j(x^i) = 0$ for $i \neq j$.

exer:3F:8 8. (a) True.

   (b) Dual basis of basis in (a) is $\varphi_0, \varphi_1, \ldots, \varphi_m$ where $\varphi_j(p) = \frac{p^{(j)}(5)}{j!}$.

exer:3F:9 9. Since $\varphi_1, \ldots, \varphi_n$ is dual basis of V' so for any $\psi \in V'$, there exists $\alpha_1, \ldots, \alpha_n \in \mathbf{F}$ so $\psi = \sum_{i=1}^n \alpha_i \varphi_i$. Note that $\left(\sum_{i=1}^n \alpha_i \varphi_i\right)(v_i) = \alpha_i$ so $\psi(v_i) = \alpha_i$. Thus, $\psi = \sum_{i=1}^n \psi(v_i)\varphi_i$.

exer:3F:10 10. We have $(S+T)'(\varphi) = \varphi \circ (S+T) = \varphi \circ S + \varphi \circ T = S'(\varphi) + T'(\varphi) = (S'+T')(\varphi)$. Thus, $(S+T)' = S' + T'$. Similarly, $(\lambda T)' = \lambda T'$.

exer:3F:11 11. If there exists such $(c_1, \ldots, c_m) \in \mathbf{F}^m$ and $(d_1, \ldots, d_n) \in \mathbf{F}^n$ then $A_{\cdot,k} = d_k \cdot C$ where C is an m-by-1 matrix so $C_i = c_i$. Hence, rank of A is the dimension of span $(A_{\cdot,1}, A_{\cdot,2}, \ldots, A_{\cdot,n})$, which is 1.

If rank of A is 1 then dimension of span $(A_{\cdot,1}, A_{\cdot,2}, \dots, A_{\cdot,n})$ is 1, which means $A_{\cdot,k} = d_k \cdot C$ for some $C \in \mathbf{F}^{m,1}$.

exer:3F:12

12. Let I be the identity map on V then the dual map $I'(\varphi) = \varphi \circ I = \varphi$ so I' is the identity map on V'.

exer:3F:13

13. (a) $(T'(\varphi_1))(x,y,z) = (\varphi_1 \circ T)(x,y,z) = \varphi_1(4x+5y+6z, 7x+8y+9z) = 4x+5y+6z$ and $(T'(\varphi_2))(x,y,z) = 7x+8y+9z$.

    (b) According to exercise 9 (3F) then since $T'(\varphi_1) \in (\mathbf{R}^3)'$ and ψ_1, ψ_2, ψ_3 is dual basis of $(\mathbf{R}^3)'$ so

$$\begin{aligned}(T'(\varphi_1)) &= (T'(\varphi_1))(1,0,0)\psi_1 + (T'(\varphi_1))(0,1,0)\psi_2 + (T'(\varphi_1))(0,0,1)\psi_3, \\ &= 4\psi_1 + 5\psi_2 + 6\psi_3.\end{aligned}$$

exer:3F:14

14. (a) We have $(T'(\varphi))(p) = (\varphi \circ T)(p) = \varphi(Tp) = \varphi(x^2p(x)+p''(x)) = (x^2p(x)+p''(x))'(4) = (2xp(x) + x^2p'(x) + p'''(x))(4) = 8p(4) + 16p'(4) + p'''(4)$.

    (b) $(T'(\varphi))(x^3) = \varphi(T(x^3)) = \varphi(x^5+6x) = \int_0^1 (x^5+6x)dx = \frac{1}{6} + 3$.

exer:3F:15

15. If $T = 0$ then $T'(\varphi) = \varphi \circ T = 0$ for any $\varphi \in W'$. Hence, $T' = 0$.

    Conversely, if $T' = 0$, let $w_1, \dots, w_m$ be basis of W. For any v, we can write $T(v) = \sum_{i=1}^m \alpha_i w_i$. We can choose $\varphi \in W$ so $\varphi(w_i) = \alpha_i$. Hence,

$$0 = (T'(\varphi))(v) = \varphi(Tv) = \sum_{i=1}^m \alpha_i \varphi(w_i) = \sum_{i=1}^m \alpha_i^2.$$

    Thus, $\alpha_i = 0$ for all $1 \le i \le m$. This follows $Tv = 0$ for all $v \in V$, i.e. $T = 0$.

exer:3F:16

16. Let S be the map that takes $T \in \mathcal{L}(V,W)$ to $T' \in \mathcal{L}(W',V')$. S is a linear map since $S(T_1+T_2) = (T_1+T_2)' = (T_1)' + (T_2)' = S(T_1) + S(T_2)$ and $S(\lambda T_1) = (\lambda T_1)' = \lambda (T_1)' = \lambda S(T_1)$.

    We prove S is injective. Indeed, if $S(T) = T' = 0$ then according to exercise 15 (3F) then $T = 0$. Thus, S is injective.

    We prove S is surjective. Let $v_1, \dots, v_n$ be basis of V and $\varphi_1, \dots, \varphi_n$ be the dual basis of V', $w_1, \dots, w_m$ be basis of W and $\psi_1, \dots, \psi_m$ be dual basis of W'. We want to determine all $A_{i,j}$ so $Tv_k = \sum_{i=1}^m A_{i,k} w_i$. For every $1 \le i \le m, 1 \le k \le n$, let $A_{i,k} = (T'(\psi_i))(v_k)$. Now we will prove that such linear map T will satisfy $T'(\psi) = \psi \circ T$ for all $\varphi \in W'$. Indeed, we find that

$$\begin{aligned}\psi_j(Tv_k) &= \psi_j\left(\sum_{i=1}^m A_{i,k} w_i\right), \\ &= \psi_j\left(\sum_{i=1}^m (T'(\psi_i))(v_k) w_i\right), \\ &= (T'(\psi_j))(v_k), \text{ (because } \psi_j(w_i) = 0 \text{ for } i \ne j)\end{aligned}$$

This is true for all $1 \le k \le n$ so for any $v \in V$ then $\psi_j(Tv) = (T'(\psi_j))(v)$ or $T'(\psi_j) = \psi_j \circ T$. This again is true for every $1 \le j \le m$ so for any $\psi \in W'$ then $T'(\psi) = \psi \circ T$, which means $S(T) = T'$. Thus, S is surjective.

In conclusion, S is an isomorphism of $\mathcal{L}(V, W)$ onto $\mathcal{L}(W', V')$.

exer:3F:17 17. For any $\varphi \in U^0$ then $\varphi(u) = 0$ for any $u \in U$, i.e. $u \in \operatorname{null} \varphi$ for any $u \in U$, i.e. $U \subset \operatorname{null} \varphi$. Thus $U^0 \subset \{\varphi \in V' : U \subset \operatorname{null} \varphi\}$. Similarly, we can finally show that $U^0 = \{\varphi \in V' : U \subset \operatorname{null} \varphi\}$.

exer:3F:18 18. $U = \{0\}$ iff $\dim U = 0$ iff $\dim U^0 = \dim V - \dim U = \dim V = \dim V'$ iff $U^0 = V'$ since U^0 is subspace of V'.

exer:3F:19 19. $U = V$ iff $\dim U = \dim V$ iff $\dim U^0 = 0$ iff $U^0 = \{0\}$.

exer:3F:20 20. If $U \subset W$ then if $\varphi \in W^0$ we will have $\varphi(w) = 0$ for any $w \in W$, which means $\varphi(u) = 0$ for any $u \in U \subset W$. Thus, $\varphi \in U^0$. Hence, $W^0 \subset U^0$.

exer:3F:21 21. Let $w_1, \ldots, w_n$ be basis of W. If U is not subspace of W then there exists $v \in V$ so $v \notin W, v \in U$. Hence, basis $w_1, \ldots, w_n$ of W can be extended to basis $w_1, \ldots, w_n, v, v_1, \ldots, v_m$ of V. Let $\varphi \in V'$ so $\varphi(w_i) = 0$ and $\varphi(v) = 1$. This follows $\varphi \in W^0$ but $\varphi \notin U^0$, a contradiction since $W^0 \subset U^0$.

exer:3F:22 22. Since $U \subset U + W$ so according to exercise 20, $(U + W)^0 \subset U^0$. Similarly, $(U + W)^0 \subset W^0$. Thus $(U + W)^0 \subset U^0 \cap W^0$.

On the other hand, if $\varphi \in U^0 \cap W^0$ then $\varphi(u) = 0$ for any $u \in U$ or $u \in W$. Thus, $\varphi(u + w) = 0$ for any $u \in U, w \in U$ so $\varphi \in (U + W)^0$. We follow $U^0 \cap W^0 \subset (U + W)^0$.

In conclusion, $U^0 \cap W^0 = (U + W)^0$.

exer:3F:23 23. For any $\varphi \in U^0, \psi \in W^0$ then $\varphi + \psi \in U^0 + W^0$ and $(\varphi + \psi)(u) = 0$ for any $u \in U \cap W$. Thus, $U^0 + W^0 \subset (U \cap W)^0$.

Since U, W are subspaces of V so $U \cap W$ is also subspace of V. Let $v_1, \ldots, v_n$ be basis of $U \cap W$, which can be extended to basis $v_1, \ldots, v_n, u_1, \ldots, u_m$ of U. Basis $v_1, \ldots, v_n$ of $U \cap W$ can also be extended to basis $v_1, \ldots, v_n, w_1, \ldots, w_k$ of W. Note that $v_1, \ldots, v_n, u_1, \ldots, u_m, w_1, \ldots, w_k$ is linearly independent so this list can be extended further to basis $v_1, \ldots, v_n, u_1, \ldots, u_m, w_1, \ldots, w_k, z_1, \ldots$ of V. Hence, for any $T \in (U \cap W)^0$, then there exists $\varphi \in U^0, \psi \in V^0$ so $\varphi(w_i) = Tw_i, \varphi(z_i) = 1$ and $\psi(u_i) = Tu_i, \varphi(z_i) = Tz_i - 1$. Hence, for any $v \in V, v = \sum_{i=1}^n \alpha_i v_i + \sum_{i=1}^m \beta_i u_i + \sum_{i=1}^k \gamma_i w_i + \sum_{i=1}^h a_i z_i$ then

$$\begin{aligned}
(\varphi + \psi)(v) &= \sum_{i=1}^m \beta_i \psi(u_i) + \sum_{i=1}^k \gamma_i \varphi(w_i) + \sum_{i=1}^h a_i (\varphi + \psi)(z_i), \\
&= \sum_{i=1}^m \beta_i T u_i + \sum_{i=1}^k \gamma_i T w_i + \sum_{i=1}^h a_i T z_i, \\
&= Tv.
\end{aligned}$$

Thus, $T = \varphi + \psi$. This follows $(U \cap W)^0 \subset U^0 + W^0$. Thus, $(U \cap W)^0 = U^0 + W^0$.

exer:3F:24 24. Done.

exer:3F:25 25. Let $S = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in U^0\}$. If $v \in U$ then obviously $v \in S$. Thus, $U \subset S$. If $v \in S$ but $v \notin U$, then there exists basis $u_1, \dots, u_n, v, v_1, \dots, v_m$ of V where $u_1, \dots, u_n$ is basis of U. Hence, there exists $\varphi \in U^0$ so $\varphi u_i = 0$ and $\varphi v = 0$. Thus, $v \notin S$, a contradiction. We find that if $v \in S$ then $v \in U$. In conclusion $U = S = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in U^0\}$.

exer:3F:26 26. Let $S = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in \Gamma\}$. If $\varphi \in \Gamma$ then $\varphi(v) = 0$ for every $v \in S$ so $\varphi \in S^0$. Thus, $\Gamma \subset S^0$.

Let $v_1, \dots, v_n$ be basis of S, which can be extended to basis $v_1, \dots, v_n, u_1, \dots, u_m$ of V. Let $\varphi_1, \dots, \varphi_n, \psi_1, \dots, \psi_m$ be dual basis of V'. Hence, for any $\varphi \in \Gamma \subset V'$, we have $\varphi = \sum_{i=1}^n \alpha_i \varphi_i + \sum_{j=1}^m \beta_j \psi_j$. Note that $\psi_1, \dots, \psi_m$ is basis of S^0 so since $v_k \in S$, we have $\psi_j(v_k) = 0$ for any $1 \le j \le m$. Thus,

$$\begin{aligned} \varphi(v_k) &= \sum_{i=1}^n \alpha_i \varphi_i(v_k) + \sum_{j=1}^m \beta_j \psi_j(v_k), \\ &= \alpha_k. \end{aligned}$$

On the other hand, since $\Gamma \subset S^0$ and $\varphi \in \Gamma, v_k \in S$ so $\varphi(v_k) = 0$. Thus, $\alpha_k = 0$ for every $1 \le k \le n$. This follows $\psi_1, \dots, \psi_m$ spans Γ. Note that $\psi_1, \dots, \psi_m$ is linearly independent. Thus, $\psi_1, \dots, \psi_m$ is basis of Γ, which leads to $\Gamma = \text{span}\,(\psi_1, \dots, \psi_m) = S^0$.

exer:3F:27 27. According to theorem 5.10.4 then $(\text{range } T)^0 = \{\lambda\varphi : \lambda \in \mathbf{R}\}$. According to exercise 25 (3F) then

$$\begin{aligned} \text{range } T &= \{p \in \mathcal{P}(\mathbf{R}) : (\lambda\varphi)(p) = 0 \text{ for any } \lambda \in \mathbf{R}\}, \\ &= \{p \in \mathcal{P}(\mathbf{R}) : \lambda p(8) = 0 \text{ for any } \lambda \in \mathbf{R}\}, \\ &= \{p \in \mathcal{P}(\mathbf{R}) : p(8) = 0\}. \end{aligned}$$

exer:3F:28 28. According to theorem 5.10.4 then $(\text{range } T)^0 = \{\lambda\varphi : \lambda \in \mathbf{F}\}$. According to exercise 25 (3F) then

$$\begin{aligned} \text{range } T &= \{v \in V : (\lambda\varphi)(v) = 0 \text{ for any } \lambda \in \mathbf{F}\}, \\ &= \{v \in V : \varphi(v) = 0\}, \\ &= \text{null } \varphi. \end{aligned}$$

exer:3F:29 29. According to theorem 5.10.6 then $(\text{null } T)^0 = \{\lambda\varphi : \lambda \in \mathbf{F}\}$. According to exercise 25 (3F) then

$$\begin{aligned} \text{null } T &= \{v \in V : (\lambda\varphi)(v) = 0 \text{ for any } \lambda \in \mathbf{F}\}, \\ &= \text{null } \varphi. \end{aligned}$$

exer:3F:30 30. Let $U = (\text{null } \varphi_1) \cap (\text{null } \varphi_2) \cap \cdots \cap (\text{null } \varphi_m)$ then U is a subspace of V. Hence, $U^0 = \{\varphi \in V' : \varphi(u) = 0 \text{ for any } u \in U\}$. Thus, we find that $\varphi_i \in U^0$ for every $1 \le i \le m$. Since $\varphi_1, \ldots, \varphi_m$ is linearly independent so $\dim U^0 \ge m$. Thus, $\dim U = \dim V - \dim U^0 \le \dim V - m$.

On the other hand, since V is finite-dimensional so let $\dim V = n \ge m$. If we let A_i be set that contains a basis of null φ_i. Then we will notice that $\dim U \ge |A_1 \cap A_2 \cap \cdots \cap A_m|$. It suffices to prove $|A_1 \cap A_2 \cap \cdots \cap A_m| \ge n - m$.

Observe that since φ_i in the linearly dependent list so range $\varphi_i = 1$. Thus, $|A_i| = \text{null } \varphi_i = \dim V - \text{range } \varphi_i = n - 1$. We prove by induction on m that $|A_1 \cap A_2 \cap \cdots \cap A_m| \ge n - m$. Indeed, it's true for $m = 1$. If it's true for $m - 1$, consider for m then we have:

$$\begin{aligned} |A_1 \cap \cdots \cap A_m| &= |A_1 \cap \cdots \cap A_{m-1}| + |A_m| - |(A_1 \cap \cdots \cap A_{m-1}) \cup A_m|, \\ &\ge n - (m-1) + (n-1) - n = n - m. \end{aligned}$$

Thus, we obtain $\dim U \ge |A_1 \cap \cdots \cap A_m| \ge n - m$. Hence, we find $\dim U = \dim V - m$.

exer:3F:31 31. According to exercise 30 (3F) then $\dim\left((\text{null } \varphi_1) \cap (\text{null } \varphi_2) \cap \cdots \cap (\text{null } \varphi_n)\right) = 0$ and $\dim\left((\text{null } \varphi_2) \cap \cdots \cap (\text{null } \varphi_n)\right) = 1$ so there exists $v_1 \in V$ so $\varphi_1(v_1) = 1$ but $\varphi_i(v_1) = 0$ for $2 \le i \le n$. Similarly, there exists $v_2, v_3, \ldots, v_n$.

Note that $v_k \notin \text{span }(v_1, \ldots, v_{k-1})$ because otherwise $1 = \varphi_k(v_k) = \sum_{i=1}^{k-1} \alpha_i \varphi_i(v_k) = 0$, a contradiction. This follows $v_1, \ldots, v_n$ is linearly independent so $v_1, \ldots, v_n$ is basis of V whose dual basis is $\varphi_1, \ldots, \varphi_n$.

exer:3F:32 32. (a) T is surjective since range $T = \text{span }(v_1, \ldots, v_n) = V$. Thus, operator T is invertible.

(b) Since T is invertible so $Tv_1, \ldots, Tv_n$ is linearly independent, which follows $\mathcal{M}(Tu_1), \ldots, \mathcal{M}(Tu_n)$ is linearly independent, i.e. columns of $\mathcal{M}(T)$ is linearly independent.

(c) Since $\dim \text{span }(\mathcal{M}(Tu_1), \ldots, \mathcal{M}(Tu_n)) = n = \dim \mathbf{F}^{n,1}$ so column of $\mathcal{M}(T)$ spans $\mathbf{F}^{n,1}$.

(d) Let $\varphi_1, \ldots, \varphi_n$ be dual basis of $u_1, \ldots, u_n$ and $\psi_1, \ldots, \psi_n$ be dual basis of $v_1, \ldots, v_n$. $T' \in \mathcal{L}(V')$ is a dual map of T with respect to these to dual bases. Since $(\mathcal{M}(T'))^t = \mathcal{M}(T)$ so i-th row of $\mathcal{M}(T)$ is i-th column of $\mathcal{M}(T')$. Since T is invertible so T' is also invertible. Hence, $T'(\varphi_1), \ldots, T'(\varphi_n)$ is linearly independent. This follows $\mathcal{M}(T'\varphi_1), \mathcal{M}(T'\varphi_2), \ldots, \mathcal{M}(T'\varphi_n)$ is linearly independent or columns of $\mathcal{M}(T')$ are linearly independent or rows of $\mathcal{M}(T)$ are linearly idependent in $\mathbf{F}^{n,1}$.

(e) Row rank of $\mathcal{M}(T)$ is n so rows of $\mathcal{M}(T)$ spans $\mathbf{F}^{n,1}$.

exer:3F:33 33. Let S be a function that takes $A \in \mathbf{F}^{m,n}$ to $A^t \in \mathbf{F}^{n,m}$ then $S(A+B) = (A+B)^t = A^t + B^t = S(A) + S(B)$ and $S(\lambda A) = (\lambda A)^t = \lambda A^t = \lambda S(A)$. Thus, S is a linear map from $\mathbf{F}^{m,n}$ onto $\mathbf{F}^{n,m}$.

S can be easily seen to be surjective, as for any $A \in \mathbf{F}^{n,m}$ then there exists $B = A^t \in \mathbf{F}^{m,n}$ so $S(B) = S(A^t) = (A^t)^t = A$.

If $S(A) = 0$ then $A = 0$. Thus, S is injective so S is invertible.

exer:3F:34 34. (a) For any $\varphi \in V'$ then $(\Lambda(v_1+v_2))(\varphi) = \varphi(v_1+v_2) = \varphi(v_1)+\varphi(v_2) = (\Lambda v_1)(\varphi)+(\Lambda v_2)(\varphi)$. Thus, $\Lambda(v_1+v_2) = \Lambda v_2 + \Lambda v_1$. Similarly, $\Lambda(\lambda v) = \lambda(\Lambda v)$. Thus, Λ is a linear map from V to V''.

(b) For any $v \in V$, we will prove that $(\Lambda \circ T)(v) = (T'' \circ \Lambda)(v)$ or $\Lambda(Tv) = T''(\Lambda v)$. In order to prove this, we show that for any $\varphi \in V'$ then $(\Lambda(Tv))(\varphi) = (T''(\Lambda v))(\varphi)$. Indeed, we have $(\Lambda(Tv))(\varphi) = \varphi(Tv)$ and since $T'' = (T')' \in \mathcal{L}(V'', V'')$ so

$$(T''(\Lambda v))(\varphi) = ((\Lambda v) \circ T')(\varphi) = (\Lambda v)(T'(\varphi)) = (T'(\varphi))(v) = (\varphi \circ T)(v) = \varphi(Tv).$$

Thus, $(T''(\Lambda v))(\varphi) = (\Lambda(Tv))(\varphi)$ for any $\varphi \in V'$, which follows $\Lambda(Tv) = T''(\Lambda v)$ for any $v \in V$, which leads to $\Lambda \circ T = T'' \circ \Lambda$.

(c) If $\Lambda v = 0$ then $(\Lambda v)(\varphi) = \varphi(v) = 0$ for any $\varphi \in V'$, which leads to $v = 0$. Thus, Λ is injective.

Let $v_1, \ldots, v_n$ be basis of V and $\varphi_1, \ldots, \varphi_n$ be dual basis of V'. For any $S \in V''$, we show that $\Lambda v = S$ where $v = \sum_{i=1}^n S(\varphi_i)v_i$. It suffices to show that. Indeed, we have $\varphi_i(v) = S(\varphi_i)$ for every $1 \le i \le n$. This follows $\varphi(v) = S(\varphi)$ for any $\varphi \in V'$. Thus, $\Lambda v = S$. Thus, Λ is surjective.

In conclusion, Λ is an isomorphism from V onto V''.

exer:3F:35 35. Denote $\varphi_i \in (\mathcal{P}(\mathbf{R}))'$ so $\varphi_i(x^i) = 1$ and $\varphi_i(x^j) = 0$ for $i \neq j$. This follows for any $\varphi \in (\mathcal{P}(\mathbf{R}))'$ then $\varphi = \sum_{i \ge 0} \varphi(x^i)\varphi_i$.

Let $e_i \in \mathbf{R}^\infty$ where $e_i = (\underbrace{0, \ldots, 0}_{i-1}, 1, 0, \ldots)$. Then any $e \in \mathbf{R}^\infty$ we have $e = \sum_{i \ge 0} \alpha_i e_i$.

Define a linear map T from $(\mathcal{P}(\mathbf{R}))'$ onto $\mathbf{F}^\infty$ so $T(\varphi_i) = e_i$. T is surjective. If $T(\varphi) = 0$ for some $\varphi = \sum_{i \ge 0} \varphi(x^i)\varphi_i$ then $T(\varphi) = \sum_{i \ge 0} \varphi(x^i)e_i = 0$. Thus, $\varphi(x^i) = 0$ for any $i \ge 0$. This follows $\varphi = 0$. Thus, T is injective. We find T is an ismorphism from $(\mathcal{P}(\mathbf{R}))'$ onto $\mathbf{F}^\infty$.

exer:3F:36 36. (a) Since range $i = U$ so according to theorem 5.10.4 then null $i' = U^0$.

(b) Since null $T = \{0\}$ so $(\text{null } T)^0 = U'$. According to theorem 5.10.6 then range $i' = (\text{null } T)^0 = U'$.

(c) Since null $i' = U^0$ so according to theorem 5.8.3 then $\tilde{i}'$ is injective and range $\tilde{i}' =$ range $i' = U'$ so $\tilde{i}'$ is surjective. Thus, $\tilde{i}'$ is an isomorphism from $V'/U^0 = V'/(\text{null } i)'$ onto U'.

exer:3F:37 37. (a) According to theorem 5.10.5, π' is injective iff π is surjective, which is true.

(b) According to theorem 5.10.6 then range $\pi' = (\text{null } \pi)^0 = U^0$.

(c) From (a) and (b) then π' is an ismorphism from $(V/U)'$ onto U^0.

6. Chapter 4: Polynomials

theo:4.17:4 **Theorem 6.0.1** (4.17) Suppose $p \in \mathcal{P}(\mathbf{R})$ is nonconstant polynomial. Then p has unique factorization (except for the order of factors) of the form

$$p(x) = c(x-\lambda_1)\cdots(x-\lambda_m)(x^2+b_1x+c_1)\cdots(x^2+b_Mx+c_M)$$

where $c, \lambda_1, \dots, \lambda_m, b_1, \dots, b_M, c_1, \dots, c_M \in \mathbf{R}$ with $b_j^2 < 4c_j$ for each j.

6.1. Exercise 4

exer:4:1 1. Done.

exer:4:2 2. No since $(-x^m+x)+x^m = x$ does not have degree of m.

exer:4:3 3. Yes.

exer:4:4 4. Pick $p = (x-\lambda_1)^{a_1}(x-\lambda_2)^{a_2}\cdots(x-\lambda_m)^{a_m}$ so $a_i \in \mathbf{Z}, a_i \geq 1$ and $\sum_{i=1}^m a_i = n$.

exer:4:5 5. Let A be a $(m+1)$-by-$(m+1)$ matrix so $A_{i,j} = z_i^{j-1}$. Define a linear map $T \in \mathcal{L}(\mathbf{R}^{m+1,1})$ so $Tx = Ax$. First, we prove that T is invertible. Indeed, if there exists $x \neq 0, x \in \mathbf{R}^{m+1,1}, x = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix}$ so $Tx = Ax = 0$ then the polynomial $a_0 + a_1x + a_2x^2 + \ldots + a_mx^m$ has $m+1$ distinct zeros $z_1, \dots, z_m, z_{m+1}$, a contradiction. Thus, we must have $x = 0$ which follows operator T is injective, which means T is invertible. Hence for any $W \in \mathbf{R}^{m+1,1}$, $W_{i,1} = w_i$ then there exists unique $x \in \mathbf{R}^{m+1,1}, x_{1,i} = \alpha_{i-1}$ so $Tx = Ax = W$. Note that $A_{i,.}x = (Ax)_{i,1} = w_i$ so polynomial $p = \alpha_0 + \alpha_1x + \alpha_2x^2 + \ldots + \alpha_mx^m$ satisfies the condition. Since the $x \in \mathbf{R}^{m+1,1}$ is uniquely determined so p is uniquely determined.

exer:4:6 6. If $p(x)$ has a zero λ then there exists $q \in \mathcal{P}(\mathbf{C})$ so $\deg q = m-1$ and $p(x) = (x-\lambda)q(x)$. Hence, $p'(x) = q(x) + (x-\lambda)q'(x)$.

   If $p(x), p'(x)$ shares a same zero, say λ then $p'(\lambda) = 0$ which follows $q(\lambda) = 0$ so $q(x) = (x-\lambda)r(x)$ so $p(x) = (x-\lambda)^2r(x)$ with $r \in \mathcal{P}(\mathbf{C})$ and $\deg r = m-2$. This follows p has at most $m-1$ distinct zeros.

exer:4:7 7. According to theorem 6.0.1 then $p(x)$ has a unique factorization

$$p(x) = c(x-\lambda_1)\cdots(x-\lambda_m)(x^2+b_1x+c_1)\cdots(x^2+b_Mx+c_M).$$

   With $p(x)$ having no real zero then each $p(x)$ has a unique factorization

$$p(x) = c(x^2+b_1x+c_1)\cdots(x^2+b_Mx+c_M)$$

   This follows degree of p is always even, a contradiction. Thus, polynomial with odd degree must have a real zero.

exer:4:8 8. There exists $q(x) \in \mathcal{P}(\mathbf{R})$ so $p(x) = (x-3)q(x) + r$ for some $r \in \mathbf{R}$. We can see right away that $r = p(3)$. We prove that $Tp = q \in \mathcal{P}(\mathbf{R})$. Indeed, $q \in \mathbf{R}^{\mathbf{R}}$ and for $x \neq 3$ then $q(x) = \frac{p(x)-r}{x-3} = \frac{p(x)-p(3)}{x-3}$. For $x = 3$ then observe that $p'(x) = q(x) + (x-3)q'(x)$ so $p'(3) = q(3)$. Thus, essentially, $Tp = q$ is a polynomial in $\mathcal{P}(\mathbf{R})$.

We have $(T(p+q))(3) = (p+q)'(3) = p'(3) + q'(3) = (T(p))(3) + (T(q))(3)$ and for $x \neq 3$ then $(T(p+q))(x) = \frac{(p+q)(x)-(p+q)(3)}{x-3} = \frac{p(x)-p(3)}{x-3} + \frac{q(x)-q(3)}{x-3} = (T(p))(x) + (T(q))(x)$. Thus, $T(p+q) = T(p) + T(q)$. Similarly, $T(\lambda p) = \lambda T(p)$. We find T is a linear map.

exer:4:9 9. If $p \in \mathcal{P}(\mathbf{C})$ then p has unique factorization of the form $p(z) = c(z-\lambda_1)\cdots(z-\lambda_m)$ where $c, \lambda_1, \ldots, \lambda_m \in \mathbf{C}$. Hence

$$\begin{aligned} q(z) &= p(z)\overline{p(\overline{z})}, \\ &= c\prod_{i=1}^{m}(z-\lambda_i)\overline{c\prod_{i=1}^{m}(\overline{z}-\lambda_i)}, \\ &= c\overline{c}\prod_{i=1}^{m}\left[(z-\lambda_i)(z-\overline{\lambda_i})\right], \\ &= |c|^2\prod_{i=1}^{m}\left(z^2 - 2(\mathrm{Re}\lambda_i)z + |\lambda_i|^2\right). \end{aligned}$$

Thus $q(z)$ is a polynomial with real coefficients.

exer:4:10 10. Let $q \in \mathcal{P}(\mathbf{R})$ so $q(x_i) = p(x_i) \in \mathbf{R}$ then according to exercise 5, q is uniquely determined in $\mathcal{P}(\mathbf{R})$.

Let $x_q = \mathcal{M}(q)$ with respect to standard basis in $\mathcal{P}(\mathbf{C})$. Repeat the construction of invertible operator T in exercise 5 (4) but this time $T \in \mathcal{L}(\mathbf{C}^{m+1,1})$. Then $Tx_p = Tx_q = (p(x_0), p(x_1), \ldots, p(x_m))^t$. Since T is injective so $x_p = x_q$, which follows $p = q$. Thus, p is a polynomial with real coefficients.

exer:4:11 11. If $\deg q \geq \deg p$ then there exists $r, s \in \mathcal{P}(\mathbf{F})$ so $q = pr + s$ with $\deg s < \deg p$. We find $q + U = pr + s + U = s + U$. And each such $s \in \mathcal{P}(\mathbf{F}), \deg s < \deg p$ can be represented as linear combination of $1, x, \ldots, x^{\deg(p)-1}$. Thus, we follow that $1+U, x+U, \ldots, x^{\deg(p)-1}+U$ spans $\mathcal{P}(\mathbf{F})/U$. And this list is linearly independent so this list is basis of $\mathcal{P}(\mathbf{F})/U$. Hence, $\dim \mathcal{P}(\mathbf{F})/U = \deg p$.

7. Chapter 5: Eigenvalues, Eigenvectors, and Invariant Subspaces

7.1. 5A: Invariant subspaces

example:eigenvalue **Example 7.1.1** Suppose $T \in \mathcal{L}(\mathbf{F}^2)$ defined by $T(w,z) = (-z,w)$ then T does not have eigenvalue if $\mathbf{F} = \mathbf{R}$. However, T has eigenvalues $\pm i$ if $\mathbf{F} = \mathbf{C}$.

theo:5.6:5A **Theorem 7.1.2** (5.6) Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. Then the following are equivalent:

(a) λ is an eigenvalue of T.

(b) $T - \lambda I$ is not injective.

(c) $T - \lambda I$ is not surjective.

(d) $T - \lambda I$ is not invertible.

This is deduced from theorem 5.6.4. theo:3.69:3D

theo:5.10:5A **Theorem 7.1.3** (5.10) Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T and $v_1, \dots, v_m$ are corresponding eigenvectors. Then $v_1, \dots, v_m$ is linearly independent.

theo:5.13:5A **Theorem 7.1.4** (5.13) Suppose V is finite-dimensional. Then each operator on V has at most $\dim V$ distinct eigenvalues.

7.2. Exercises 5A

exer:5A:1 1. (a) For any $u \in U$ then $Tu = 0 \in U$ so U is invariant under T.

   (b) For any $u \in U$ then $Tu \in \operatorname{range} T \subset U$ so U is invariant under T.

exer:5A:2 2. For every $u \in \operatorname{null} S$ then $STu = T(Su) = T(0) = 0$ so $Tu \in \operatorname{null} S$. Thus, $\operatorname{null} S$ is invariant under T.

exer:5A:3 3. For every $u \in \operatorname{range} S$ then there exists $v \in V$ so $Sv = u$. Hence, $Tu = T(Sv) = S(Tv) \in \operatorname{range} S$ so $\operatorname{range} S$ is invariant under T.

exer:5A:4 4. For any $u_1 \in U_1, u_2 \in U_2, \dots, u_m \in U_m$ then $T(u_1 + \ldots + u_m) = Tu_1 + \ldots + Tu_m \in U_1 + \ldots + U_m$. Thus, $U_1 + \ldots + U_m$ is invariant under T.

exer:5A:5 5. Let $U_1, \dots, U_m$ be invariant subspaces of V under T. Then for any $v \in S = U_1 \cap \ldots \cap U_m$ then $Tv \in V$ so S is invariant under T.

exer:5A:6 6. If $U \neq \{0\}$ then there exists basis $u_1, \dots, u_m$ of U. If there exists $v \in V, v \notin U$ then we can choose operator $T \in \mathcal{L}(V)$ so $Tu_1 = v$, which follows U is not invariant under T, a contradiction. Thus, if $v \in V$ then $v \in U$, i.e. $U = V$.

exer:5A:7 7. If $T(x,y) = (-3y,x) = \lambda(x,y)$ for $(x,y) \neq (0,0)$ then $-3y = \lambda x, x = \lambda y$ which follows $-3y = \lambda^2 y$. If $y = 0$ then $x = 0$, a contradiction. Hence, $y \neq 0$ which means $-3 = \lambda^2$. Thus, T does not have any eigenvalue.

exer:5A:8 8. If $T(w,z) = (z,w) = \lambda(w,z)$ for $(w,z) \neq (0,0)$ then $z = \lambda w, w = \lambda z$ so $z = \lambda^2$. Similarly, since $z \neq 0$ so $\lambda^2 = 1$ which follows $\lambda = \pm 1$. If $\lambda = 1$ then eigenvector is (z,z), if $\lambda = -1$ then eigenvector is $(z,-z)$.

exer:5A:9 9. If $T(z_1,z_2,z_3) = \lambda(z_1,z_2,z_3) = (2z_1,0,5z_3)$ for $(z_1,z_2,z_3) \neq 0$ then $\lambda z_1 = 2z_2, \lambda z_2 = 0, \lambda z_3 = 5z_3$. If $\lambda = 0$ then $z_2 = z_3 = 0$ so eigenvector is $(z,0,0)$. If $\lambda ne 0$ then $z_1 = z_2 = 0$ and $\lambda z_3 = 5z_3$ so $\lambda = 5$. Respective eigenvector is $(0,0,z)$.

exer:5A:10 10. (a) If $T(x_1,\ldots,x_n) = \lambda(x_1,\ldots,x_n) = (x_1,2x_2,3x_3,\ldots nx_n)$ for $(x_1,\ldots,x_n) \neq (0,\ldots,0)$ then $(\lambda - i)x_i = 0$ for $i = 1,\ldots,n$. This follows $\lambda \in \{1,\ldots,n\}$. If $\lambda = i$ then eigenvector is $\alpha \cdot e_i$ where $e_i \in \mathbf{F}^n$ so e_i's coordinates are 0 except the i-th coordinate is 1.

(b) Let U be an invariant subspace of V under T then for any $u = (x_1,\ldots,x_n) \in U$ then $v = (x_1,2x_2,\ldots,nx_n) \in U$. With only these two vectors in U, we can choose certain $\alpha,\beta \in \mathbf{F}$ to 1-st coordinate of $w = \alpha u + \beta v$ is 0. We also have $Tw \in U$ and from w, Tw, we can obtain a new vector in U so its 1st and 2nd coordinates are 0. Keep going like that, we can eventually obtain that $e_i = (\underbrace{0,\ldots,0}_{i-1},1,0,\ldots 0) \in U$ if $x_i \neq 0$. This follows subspace U is spanned by list of e_i so $x_i \neq 0$ for some $(x_1,\ldots,x_n) \in U$. Thus, any subset of the standard basis of $\mathbf{F}^n$ is a spanning list of an invariant subspace of V.

exer:5A:11 11. If $Tp = \lambda p = p'$ for $p \neq 0$. If $\lambda = 0$ then $p' = 0$ so respective eigenvector p is a constant. If $\lambda \neq 0$ then since $\lambda p \neq p'$ for all $p \in \mathcal{P}(\mathbf{R})$ so doesn't exists eigenvector in this case. Thus, the only answer is $\lambda = 0$ as eigenvalue and $p = c \in \mathbf{R}$ as eigenvector.

exer:5A:12 12. If $(Tp)x = \lambda p(x) = xp'(x)$ for some $p(x) \neq 0$. Let $p = a_0 + a_1x + \ldots, a_4x^4$ then $p'(x) = a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3$. Thus, this follows $\lambda a_i = ia_i$ for $i = 0,\ldots,4$. Hence, $\lambda \in \{0,1,2,3,4\}$ so that $(a_0,\ldots,a_4) \neq (0,\ldots,0)$. If $\lambda = i$ then $p = ax^i$ is can eigenvector for any $a \in \mathbf{R}$.

exer:5A:13 13. Since V is finite-dimensional so according to theorem 7.1.4, there are at most $\dim V$ different eigenvalues of T, in other words, at most $\dim V$ $\alpha \in \mathbf{F}$ so $T - \alpha I$ is not invertible according to theorem 7.1.2. Thus, we can clearly choose α so $|\alpha - \lambda| < \frac{1}{1000}$ and $T - \alpha I$ is invertible.

exer:5A:14 14. If $P(u+w) = \lambda(u+w) = u$ for $u + w \neq 0, u \in U, w \in W$ then $(1-\lambda)u = \lambda w$. If $\lambda \neq 0,1$ then $w \in U$, a contradiction since $V = U \oplus W$. If $\lambda = 0$ then $u = 0$ so eigenvector is any $w \in W$. If $\lambda = 1$ then $w = 0$ so eigenvector is any $u \in W$.

exer:5A:15 15. (a) If λ is an eigenvalue of T then there exists $v \in V, v \neq 0$ so $Tv = \lambda v$. Since S is invertible so there exists $u \in V$ so $Su = v$ which follows $S^{-1}v = u$. We have $S^{-1}TSu = S^{-1}T(Su) = S^{-1}Tv = S^{-1}(\lambda v) = \lambda u$. Thus, λ is also eigenvalue of $S^{-1}TS$.

If λ is an eigenvalue of $S^{-1}TS$ then there exists $v \in V, v \neq 0$ so $S^{-1}TSv = \lambda v$. or $S^{-1}(TSv) = \lambda S^{-1}(Sv)$. Since S^{-1} is invertible so $T(Sv) = \lambda Sv$. Thus, λ is also eigenvalue of T.

(b) If v is eigenvalue of T then $S^{-1}v$ is eigenvector of $S^{-1}TS$ corresponding to λ.

exer:5A:16 16. Let $v_1, \ldots, v_n$ be basis of V and for $1 \leq k \leq n$, let $Tv_k = \sum_{i=1}^n \alpha_{i,k} v_i$ so $\alpha_{i,j} \in \mathbf{R}$ according to problem's condition. If λ is eigenvalue of T with eigenvector $v = \sum_{i=1}^m \beta_i v_i$ with $\beta_i \in \mathbf{C}$. We show that $\overline{v} = \sum_{i=1}^n \overline{\beta_i} v_i$ is the eigenvector corresponding to eigenvalue $\overline{\lambda}$ of T, i.e. $T\overline{v} = \lambda \cdot \overline{v}$.

Indeed, from $Tv = \lambda v$, where $Tv = \sum_{k=1}^n \beta_k T v_k = \sum_{i=1}^n \beta_k \sum_{i=1}^n \alpha_{i,k} v_i$, we follow for any $i = 1, \ldots, n$ then $\sum_{k=1}^n \beta_k \alpha_{i,k} = \lambda \beta_i$ so $\overline{\sum_{k=1}^n \beta_k \alpha_{i,k}} = \overline{\lambda \beta_i}$ or $\sum_{k=1}^n \overline{\beta_k} \alpha_{i,k} = \overline{\lambda} \cdot \overline{\beta_i}$. Thus,

$$T\overline{v} = \sum_{i=1}^n \left(\sum_{k=1}^n \overline{\beta_k} \alpha_{i,k} \right) v_i = \sum_{i=1}^n (\overline{\lambda} \cdot \overline{\beta_i} v_i) = \overline{\lambda} \cdot \overline{v}.$$

Thus, $\overline{\lambda}$ is also an eigenvalue of T.

exer:5A:17 17. $T(x_1, x_2, x_3, x_4) = (x_2, x_3, x_4, -x_1)$.

exer:5A:18 18. If $T(z_1, z_2, \ldots) = (0, z_1, \ldots) = \lambda(z_1, z_2, \ldots)$ for some $(z_1, z_2, \ldots) \neq (0, 0, \ldots)$ then $\lambda z_1 = 0$ and $\lambda z_{i+1} = z_i$ for all $i \geq 1$. In both two cases $\lambda = 0$ and $\lambda \neq 0$, we all obtain $z_i = 0$ for all $i \geq 1$, a contradiction. Thus, T does not have eigenvalue.

exer:5A:19 19. If $T(x_1, \ldots, x_n) = \lambda(x_1, \ldots, x_n) = (x_1 + \ldots + x_n, \ldots, x_1 + \ldots + x_n)$ for some $(x_1, \ldots, x_n) \neq (0, \ldots, 0)$ then $\lambda x_i = x_1 + \ldots + x_n$ for all $1 \leq i \leq n$. Adding all n equations we find $\lambda(x_1 + \ldots + x_n) = n(x_1 + \ldots + x_n)$. If $\lambda \neq n$ then $x_1 + \ldots + x_n = 0$, which follows $x_i = 0$ for all $1 \leq i \leq n$, a contradiction. Thus, $\lambda = n$, which leads to $x_i = x_j$ for any $1 \leq i < j \leq n$. This follows respective eigenvector is $\alpha \cdot (1, 1, \ldots, 1)$ for any $\alpha \in \mathbf{F}$.

exer:5A:20 20. If $T(z_1, z_2, \ldots) = \lambda(z_1, z_2, \ldots) = (z_2, z_3, \ldots)$ for some $(z_1, z_2, \ldots) \neq (0, 0, \ldots)$ then $\lambda z_i = z_{i+1}$ for all $i \geq 1$. For any eigenvalue λ then corresponding eigenvector is $\alpha(1, \lambda, \lambda^2, \ldots)$ for any $\alpha \in \mathbf{F}$.

exer:5A:21 21. If λ is eigenvalue of T then there exists $v \in V, v \neq 0$ so $Tv = \lambda v$. Hence, $T^{-1}(v) = T^{-1}\left(\frac{1}{\lambda} Tv\right) = \frac{1}{\lambda} v$ or $\frac{1}{\lambda}$ is eigenvalue of T^{-1}. Similarly if $\frac{1}{\lambda}$ is eigenvalue of T^{-1} then λ is eigenvalue of T. We also find that T and T^{-1} have the same eigenvectors.

exer:5A:22 22. If $w = -v$ then -3 is eigenvalue of T. If $w + v \neq 0$ then $T(v + w) = 3(v + w)$ so 3 is eigenvalue of T.

exer:5A:23 23. If λ is eigenvalue of ST with $STv = \lambda v$ for $v \in V, v \neq 0$. Then $TS(Tv) = T(STv) = T(\lambda v) = \lambda Tv$. Since $Tv \neq 0$ so λ is also eigenvalue of TS. Similarly, if λ is eigenvalue of TS then λ is also eigenvalue of ST.

exer:5A:24 24. (a) Let $x = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ then $Tx = Ax = x$ so 1 is eigenvalue of T.

(b) Note that matrix A is precisely $\mathcal{M}(T)$ and according to theorem 5.10.7, dim range $(T - I)$ is equal to column rank of $A - I$. Since sum of all entries in each column of A is 1 so sum of the entries in each column of $A - I$ is 0. Let $e_i \in \mathbf{F}^{n,1}$ be the i-th row of $A - I$ then $-e_n = e_1 + \ldots + e_{n-1}$ so row rank of $A - I$ is at most $n - 1$, which means column rank of $A - I$ is at most $n - 1$ or dim range $(T - I) \leq n - 1$ so $T - I$ is not injective. According to theorem 7.1.2, we follow 1 is an eigenvalue of T.

exer:5A:25 25. Let $Tu = \lambda_1 u, Tv = \lambda_2 v$ and $T(u + v) = \lambda_3(u + v)$ then $\lambda_1 u + \lambda_2 v = \lambda_3(u + v)$ or $(\lambda_1 - \lambda_3)u = (\lambda_3 - \lambda_2)v$. If $\lambda_1 = \lambda_3$ then $\lambda_2 = \lambda_3 = \lambda_1$, as desired. If $\lambda_1 \neq \lambda_3$ then $\lambda_3 \neq \lambda_2$ which follows $u = \alpha v$ for $\alpha \in \mathbf{F}, \alpha \neq 0$. Hence, $Tu = \alpha Tv$ or $\lambda_1 u = \alpha \lambda_2 v$ or $\alpha v(\lambda_1 - \lambda_2) = 0$. Thus, $\lambda_1 = \lambda_2$ in all cases.

exer:5A:26 26. According to exercise 25 (5A) we can follows all nonzero vectors in V corresponds to same eigenvalue λ, i.e. $Tv = \lambda v$ for all $v \in V$. Thus, $T = \lambda I$.

exer:5A:27 27. Let $v_1, \ldots, v_n$ be basis of V and $Tv_1 = \sum_{i=1}^n \alpha_i v_i$. Consider $n - 1$-dimensional invariant subspace $U_i = \text{span}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$ $(i \neq 1)$ then since $v_1 \in U_i$ we follow $Tv_1 \in U_i$ which leads to $\alpha_i = 0$. This is true for all $i \neq 1$ so eventually we obtain $Tv_1 = \alpha_1 v_i$. Similarly, we can find $Tv_k = \alpha_k v_k$ for $k = 1, \ldots, n$.

Now consider $n-1$-dimensional invariant subspace $S_1 = \text{span}\ (v_1 - v_2, v_1 - v_3, \ldots, v_1 - v_n)$ then since $v_1 - v_2 \in S_1$ so $\alpha_1 v_1 - \alpha_2 v_2 = T(v_1 - v_2) = \sum_{i=2}^n \beta_i(v_1 - v_i)$. Note that $v_1, \ldots, v_n$ is linearly independent so we must have $\beta_i = 0$ for all $i \geq 3$. We obtain $\alpha_1 v_1 - \alpha_2 v_2 = \beta_2(v_1 - v_2)$, which brings us back to exercise 25 (5A) and obtain $\alpha_1 = \alpha_2$. Similarly, we can obtain $\alpha_i = \alpha_j = \lambda$ for all $1 \leq i < j \leq n$, which leads to $Tv = \lambda v$ for all $v \in V$. Thus, $T = \lambda I$.

exer:5A:28 28. Similarly to exercise 27 (5A), let $v_1, \ldots, v_n$ $(n \geq 3)$ be basis of V then from $v_1 \in \text{span}\ (v_1, v_2)$ and $v_1 \in \text{span}\ (v_1, v_3)$, we deduce $Tv_1 = \alpha_1 v_1$. Similarly, $Tv_i = \alpha_i v_i$ for all $i = 1, \ldots, n$.

Consider $v_1 - v_2$ in 2-dimensional invariant subspace span $(v_1 - v_2, v_1 - v_3)$ so $\alpha_1 v_i - \alpha_2 v_2 = T(v_1 - v_2) = \beta_1(v_1 - v_2) + \beta_2(v_1 - v_3)$. Since v_1, v_2, v_3 is linearly independent so $\beta_2 = 0$ or $\alpha_1 v_1 - \alpha_2 v_2 = \beta_1(v_1 - v_2)$, which according to exercise 25 (5A) then $\alpha_1 = \alpha_2$. Similarly, we can show that $\alpha_i = \alpha_j = \alpha$ for all $1 \leq i < j \leq n$. This follows $Tv = \alpha v$ for every $v \in V$ or $T = \alpha I$.

exer:5A:29 29. Let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be all distinct eigenvalues of T and $v_1, \ldots, v_m$ be corresponding eigenvectors of T then according to theorem 7.1.3, $v_1, \ldots, v_m$ is linearly independent. If $\lambda_i \neq 0$ for all $i = 1, \ldots, m$ then $Tv_i = \lambda_i v_i \in \text{range}\ T$ so $v_i \in \text{range}\ T$ for all $i = 1, \ldots, m$. This follows $m \leq \dim \text{range}\ T = k$. If there exists $\lambda_1 = 0$ then $\lambda_i \neq 0$ for all $i \geq 2$, which leads to $v_i \in \text{range}\ T$ for all $i \geq 2$ and we obtain $m - 1 \leq k$ in this case. In conclusion, $m \leq k + 1$.

exer:5A:30 30. Let x_1, x_2, x_3 be eigenvectors corresponding to $-4, 5, \sqrt{7}$ then x_1, x_2, x_3 is linearly independent so x_1, x_2, x_3 is basis of $\mathbf{R}^3$. Say $x = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3$ then $Tx - 9x = -13\alpha_1 x_1 - 4\alpha_2 x_2 + (\sqrt{7} - 9)\alpha_3 x_3$. We write $(-4, 5, \sqrt{7})$ as linear combination of x_1, x_2, x_3 then from that to find α_i so $Tx - 9x = (-4, 5, \sqrt{7})$. Thus, such $x \in \mathbf{R}^3$ so $Tx - 9x = (-4, 5, \sqrt{7})$ exists.

exer:5A:31 31. If $v_1, \ldots, v_m$ is linearly independent, since V is finite-dimensional, we can construct $T \in \mathcal{L}(V)$ so $Tv_i = \lambda_i v_i$ for $i = 1, \ldots, m$. Hence, T has eigenvectors $v_1, \ldots, v_m$ corresponding to $\lambda_1, \ldots, \lambda_m$.

exer:5A:32 32. The hint says it all.

exer:5A:33 33. Because $T(v + \text{range } T) = Tv + \text{range } T = 0 + \text{range } T$.

exer:5A:34 34. If null T is injective but $(\text{null } T) \cap (\text{range } T) \neq \{0\}$, i.e. there exists $v \notin \text{null } T$ but $Tv \in \text{null } T$, i.e. $Tv \in (\text{null } T) \cap (\text{range } T)$. Hence, $(T/\text{null } T)(v + \text{null } T) = \text{null } T$. Since $v \notin \text{null } T$ so we find $T/\text{null } T$ is not injective, a contradiction. Thus, we must have $(\text{null } T) \cap (\text{range } T) = \{0\}$.

Conversely, if $(\text{null } T) \cap (\text{range } T) = \{0\}$. If $T/(\text{null } T)$ is not injective then there exists $v \notin \text{null } T$ so $Tv \in \text{null } T$. This follows $Tv \neq 0$ and $Tv \in (\text{null } T) \cap (\text{range } T)$, a contradiction. Thus, $T/\text{null } T$ is injective.

exer:5A:35 35. If λ is eigenvalue of T/U then there exists $v \in V, v \notin U$ so $(T/U)(v+U) = Tv+U = \lambda v+U$. This follows $Tv - \lambda v \in U$.

Consider $(T - \lambda I)|_U \in \mathcal{L}(U)$, if $(T - \lambda I)|_U$ is not injective then there exists $u \in U, u \neq 0$ so $(T - \lambda I)u = 0$ or $Tu = \lambda u$ so λ is eigenvalue of T. If $(T - \lambda I)|_U$ is injective then $(T - \lambda I)|_U$ is invertible according to theorem 5.6.4 (3D) as V is finite-dimensional. Hence for $Tv - \lambda v = u \in U$ there exists $w \in U$ so $(T - \lambda I)w = u = Tv - \lambda v$ which follows $T(v - w) = \lambda(v - w)$. Since $v - w \neq 0$ as $v \notin U$ so we also find λ as eigenvalue of T.

In conclusion, each eigenvalue of T/U is an eigenvalue of T.

exer:5A:36 36. Pick $V = \mathcal{P}(\mathbf{R})$ and $\lambda \in \mathbf{R}, \lambda \neq 0$. Let $T \in \mathcal{L}(V)$ so $T(1) = 0, T(x) = x^3 + \lambda x^2, Tx^2 = x^4 + \lambda x$ and $Tx^k = x^{7+k}$ for all $k \geq 3$. Pick subspace of V as $U = \text{span }(x^3, x^4, \ldots)$. Hence, with $v = x + x^2$ we have $Tv - \lambda v = x^3 + x^4 \in U$ so $Tv + U = \lambda v + U$, i.e. λ is an eigenvalue of T/U.

However, $Tv \neq \lambda v$ for all $v \in V$. Indeed, if $\deg v \geq 3$ then $\deg Tv \geq 7 + \deg v$, which leads to $Tv \neq \lambda v$. If $\deg v \leq 2$, let $v = a_0 + a_1 x + a_2 x^2$ then $Tv = a_1(x^3 + \lambda x^2) + a_2(x^4 + \lambda x)$ and $\lambda v = \lambda(a_0 + a_1 x + a_2 x^2)$. In this case, $Tv = \lambda v$ when $a_1 = a_2 = 0$ or $Tv = 0 = \lambda a_0$ and $v = a_0 \neq 0$, a contradiction.

Thus, in this example, λ is eigenvalue of T/U but not of T.

7.3. 5B: Eigenvectors and Upper-Triangular Matrices

theo:5.21:5B **Theorem 7.3.1** (5.21) Every operator on a finite-dimensional, nonzero, complex vector space has an eigenvalue.

theo:5.27:5B **Theorem 7.3.2** (5.27) Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some basis of V.

theo:5.30:5B **Theorem 7.3.3** (5.30) Suppose $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V. Then T is invertible if and only if all the entries on the diagonal of that upper-triangular matrix are nonzero.

In general, $T - \lambda I$ is invertible iff all entries on the diagonal of that upper-triangular matrix are different from λ.

theo:5.32:5B **Theorem 7.3.4** (5.32) Suppose $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V. Then the eigenvalues of T are precisely the entries on the diagonal of that upper-triangular matrix.

Following theorems are not from the book. I found these from a blog post of tastymath75025 (AoPS).

theo:1:5B **Theorem 7.3.5** Every operator T on a finite-dimensional real vector space V has invariant subspace U of dimension 1 or 2.

theo:2:5B **Theorem 7.3.6** Every operator T on a odd-dimensional real vector space has an eigenvalue.

From the above we follow

Proposition 7.3.7

If V is a real vector space with even dimension, then for any operator T on V, every invariant subspace of V under T has even dimension.

7.4. Exercises 5B

exer:5B:1 1. (a) We have $(I-T)(I+T+T^2+\ldots+T^{n-1}) = I-T^n = I$ and $(I+T+\ldots+T^{n-1})(I-T) = I$ so $I - T$ is invertible and $(I-T)^{-1} = I + T + T^2 + \ldots + T^{n-1}$.

   (b) Looks like $1 - x^n = (1-x)(1+x+\ldots+x^{n-1})$.

exer:5B:2 2. If λ is eigenvalue of T then let v be the eigenvector corresponding to λ. We have $((x-2)(x-3)(x-4))(T))v = 0$ or $(\lambda-2)(\lambda-3)(\lambda-4) = 0$. Thus, $\lambda = 2$ or $\lambda = 3$ or $\lambda = 4$.

exer:5B:3 3. We have $T^2 - I = (T+I)(T-I)$ so for any $v \in V, v \neq 0$ then $(T+I)((T-I)v) = 0$. If there exists v so $(T-I)v = w \neq 0$ then $(T+I)w = 0$ which follows -1 is eigenvalue of T, a contradiction. Thus, $(T-I)v = 0$ for all $v \in V$, which follows $T = I$.

exer:5B:4 4. We have $0 = P^2 - P = P(P - I)$ so for every $v \in V$ then $P(P-I)v = 0$ which follows $(P-I)v \in \text{null } P$. Hence, any $v \in V$ can be written as $v = Pv - (P-I)v$ where $Pv \in \text{range } P, (P-I)v \in \text{null } P$. We also note that $(\text{null } P) \cap (\text{range } P) = \{0\}$, otherwise there exists $u \notin \text{null } P$ but $Pu \in \text{null } P$ which leads to $P^2 u = 0 \neq Pu$, a contradiction. Thus, $V = \text{null } P \oplus \text{range } P$.

exer:5B:5 5. Note that $(STS^{-1})^k = (STS^{-1})(STS^{-1}) \cdots (STS^{-1}) = ST^kS^{-1}$ so if $p = a_0 + a_1 x + a_2 x^2 + \ldots + a_m x^m$ then

$$\begin{aligned} p(STS^{-1}) &= a_0 I + a_1 STS^{-1} + \ldots + a_m (STS^{-1})^m, \\ &= a_0 I + a_1 STS^{-1} + \ldots + a_m ST^m S^{-m}, \\ &= S(a_0 + a_1 T + \ldots + a_m T^m) S^{-1}, \\ &= Sp(T)S^{-1}. \end{aligned}$$

exer:5B:6 6. Since U is invariant under T so for any $u \in U$ then $T^n u \in U$ for all $i \geq 0$. This follows $p(T)u \in U$ for every $u \in U$, i.e. U is invariant under $p(T)$ for any $p \in \mathcal{P}(\mathbf{F})$.

exer:5B:7 7. Since 9 is eigenvalue of T^2 so there exists $v \in V, v \neq 0$ so $((z^2 - 9)(T))v = (T^2 - 9I)v = 0$ or $(((z-3)(z+3))(T))v = (T-3I)(T+3I)v = 0$. If $(T+3I)v = 0$ then -3 is eigenvalue of T. If $(T+3I)v = u \neq 0$ then $(T-3I)u = 0$ which follows 3 is eigenvalue of T.

exer:5B:8 8. Imagine T as rotation $45°$ clockwise in the plane $\mathbf{R}^2$. This will give $T^4 = -I$. Note that $T(x,y) = (x \cos 45° - y \sin 45°, x \sin 45° + y \cos 45°)$.

exer:5B:9 9. If λ is a zero of p then $p = (z - \lambda) q(z)$. Hence $0 = p(T)v = (T - \lambda I) q(T) v$. Since $\deg q < \deg p$ so $q(T)v = u \neq 0$ which follows $(T - \lambda I)u = 0$ or λ is eigenvalue of T.

exer:5B:10 10. Let $p = a_0 + a_1 z + \ldots + a_m z^m$ then from $Tv = \lambda v$ we have

$$\begin{aligned} p(T)v &= (a_0 + a_1 T + \ldots + a_m T^m) v, \\ &= a_0 v + a_1 T v + \ldots + a_m T^m v, \\ &= a_0 v + a_1 \lambda v + \ldots + a_m \lambda^m v, \\ &= p(\lambda) v. \end{aligned}$$

exer:5B:11 11. If $\alpha = p(\lambda)$ for some eigenvalue λ of T then from exercise 10 (5B) we have α is eigvalue of $p(T)$. Conversely, if α is eigenvalue of $p(T)$ then there exists $v \in V, v \neq 0$ so $p(T)v = \alpha v$ or $(p(T) - \alpha I)v = 0$. Since $\mathbf{F} = \mathbf{C}$ so we can factorise $p - \alpha = c(x - \lambda_1) \cdots (x - \lambda_m)$ which follows $0 = (p(T) - \alpha I)v = c(T - \lambda_1 I) \cdots (T - \lambda_m I)v$. Hence, there exists $1 \leq j \leq m$ so λ_j is eigenvalue of T. Hence $p(T)v = p(\lambda_j)v = \alpha v$ so $\alpha = p(\lambda_j)$.

See MSE, the problem is still true if T is only triangularizable on any finite vector space over a field F

exer:5B:12 12. Back to exercise 8 (5B), there exists $T \in \mathcal{L}(\mathbf{R}^2)$ so $T^4 = -I$, i.e. -1 is eigenvalue of $(x^4)(T)$. However, $-1 \neq p(\lambda) = \lambda^4$ for any $\lambda \in \mathbf{R}$.

exer:5B:13 13. If subspace $U \neq \{0\}$ is invariant under T. U is finite-dimensional then $T|_U \in \mathcal{L}(U)$ so according to theorem 7.3.1, there exists eigenvalue λ of $T|_U$, which follows λ is eigenvalue of T, a contradiction. Thus U is either $\{0\}$ or infinite-dimensional.

exer:5B:14 14. Pick T so $\mathcal{M}(T) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

exer:5B:15 15. Pick T so $\mathcal{M}(T) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

exer:5B:16 16. If $\dim V = n$ then consider linear map $S : \mathcal{P}_n(\mathbf{C}) \to V$ so $Sp = (p(T))v$. Since $\dim \mathcal{P}_n(\mathbf{C}) = n + 1 > n = \dim V$ so according to theorem 5.3.6, S is not injective. Hence, there exists $p \in \mathcal{P}_n(\mathbf{C}), p \neq 0$ so $Sp = (p(T))v = 0$, which leads to existence of eigenvalue of T.

exer:5B:17 17. If $\dim V = n$ then consider linear map $S : \mathcal{P}_{n^2}(\mathbf{C}) \to \mathcal{L}(V)$ so $Sp = p(T)$. Since $\dim \mathcal{P}_{n^2}(\mathbf{C}) = n^2 + 1 > n^2 = \dim \mathcal{L}(V)$ so according to theorem 5.3.6, S is not injective, i.e. there exists $p \in \mathcal{P}_{n^2}(\mathbf{C}), p \neq 0$ so $Sp = p(T) = 0$. Pick $v \neq 0, v \in V$ then $p(T)v = 0$ which follows existence of eigenvalue of T.

exer:5B:18 18. Since V is finite-dimensional complex vector space so there exists λ so $T - \lambda I$ is invertible or is not invertible. If λ is not eigenvalue of T then $T - \lambda I$ is invertible so $\dim \text{range } (T - \lambda I) = \dim v$. If λ is eigenvalue of T then $T - \lambda I$ is not invertible so $\dim \text{range } (T - \lambda I) \leq \dim v - 1$. Thus, $f(\lambda)$ is not a continuous function.

exer:5B:19 19. If $\mathbf{F} = \mathbf{C}$ then there exists eigenvalue λ of T with eigenvector $v \neq 0$. Hence, for any $p \in \mathcal{P}(\mathbf{C})$ then $p(T)v = p(\lambda)v$ according to exercise 10 (5B). On the other hand, we can choose $S \in \mathcal{L}(V)$ so $Sv \notin \text{span } v$ as $\dim V > 1$. Hence, $S \notin \{p(T) : p \in \mathcal{P}(\mathbf{C})\}$ so $\{p(T) : p \in \mathcal{P}(\mathbf{C})\} \neq \mathcal{L}(V)$.

If $\mathbf{F} = \mathbf{R}$, if there exists eigenvalue of T then it's similar to previous case. If T has no eigenvalues, i.e. T has no invariant subspace with dimension 1 then according to theorem **??**, T has invariant subspace U with dimension 2. Hence, for any $v \in U$ then $p(T)v \in U$ for any $p \in \mathcal{P}(\mathbf{R})$. If $\dim V > 2$ then we can choose $S \in \mathcal{L}(V)$ so $Su \notin U$ for $u \in U$, which follows $S \notin \{p(T) : p \in \mathcal{P}(\mathbf{R})\}$. Thus, $\{p(T) : p \in \mathcal{P}(\mathbf{R})\} \neq \mathcal{L}(V)$.

If $\dim V = 2$ then $U = V$. According to proof of theorem 7.3.5 and the fact that V has no eigenvalue, we can choose $v \in V$ so $V = U = \text{span } (v, Tv)$ for some $v \in V, v \neq 0$. We choose $S \in \mathcal{L}(V)$ so $Sv = \lambda v$ and $S(Tv) = \beta v$ for $\beta, \lambda \in \mathbf{R}; \beta, \lambda \neq 0$. If $S \in \{p(T) : p \in \mathcal{P}(\mathbf{R})\}$ then $S = a_1 I + a_2 T$. Note that $a_2 \neq 0$, otherwise $\beta v = S(Tv) = (a_1 I)(Tv) = a_1 Tv$, a contradiction since v, Tv is linearly independent. Thus, from $\lambda v = Sv = (a_1 I + a_2 T)(v)$ we obtain $Tv = \frac{\lambda - a_1}{a_2} v$, a contradiction since V has no eigenvalue. Thus, $S \notin \{p(T) : p \in \mathcal{P}(\mathbf{R})\}$.

Thus, in all cases, we have $\{p(T) : p \in \mathcal{P}(\mathbf{F})\} \neq \mathcal{L}(V)$.

exer:5B:20 20. It is a consequence of theorem 7.3.2. Since V is finite-dimensional complex vector space so there exists a basis $v_1, \ldots, v_n$ of V so $\mathcal{M}(T)$ with respect to this basis is upper-triangular matrix. Hence, for any $1 \leq i \leq n$, i-dimensional subspace $U_i = \text{span } (v_1, \ldots, v_i)$ of V is invariant under T.

7.5. 5C: Eigenspaces and Diagonal Matrices

theo:5.38:5C **Theorem 7.5.1** (5.38) Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Suppose also that $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T. Then $E(\lambda_1, T) + \dots + E(\lambda_m, T)$ is a direct sum. Furthermore,

$$\dim E(\lambda_1, T) + \dots + \dim E(\lambda_m, T) \le \dim V.$$

theo:5.41:5C **Theorem 7.5.2** (5.41) Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_m$ denote the distinct eigenvalues of T. Then the following are equivalent:

(a) T is diagonalizable;

(b) V has basis consisting of eigenvectors of T;

(c) there exists 1-dimensional subspace $U_1, \dots, U_n$ of V, each invariant under T so $V = U_1 \oplus \cdots \oplus U_n$.

(d) $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$.

(e) $\dim V = \dim E(\lambda_1, T) + \dots + \dim E(\lambda_m, T)$.

theo:5.44:5C **Theorem 7.5.3** (5.44) If $T \in \mathcal{L}(V)$ has $\dim V$ distinct eigenvalues then T is diagonalizable.

7.6. Exercises 5C

exer:5C:1 1. Since T is diagonalizable so V has basis $v_1, \dots, v_n$ consisting of eigenvectors of T. Let $Tv_i = \lambda_i v_i$. List of v_j so $\lambda_j = 0$ is basis of null space of T. For all v_j so $\lambda_j \neq 0$ then we find $Tv_j = \lambda_j v_j$ spans range of T, hence v_j is basis of range of T. Thus, $V = \text{null } T \oplus \text{range } T$.

exer:5C:2 2. The converse is not true. For example, if v_1, v_2, v_3 is basis of V, let $T \in \mathcal{L}(V)$ so $Tv_1 = 0, Tv_2 = v_3, Tv_3 = -v_2$ then span (v_1) is null space of T and span (v_2, v_3) is range of T. Hence, $V = \text{null } T \oplus \text{range } T$ but T is not diagonalizable as 0 is the only eigenvalue of T with $E(0, T) = \text{span } (v_1)$.

exer:5C:3 3. If (a) is true then (b) is also true. Let $u_1, \dots, u_n$ be basis of null T which can be extended to basis $u_1, \dots, u_n, v_1, \dots, v_m$ of V.

If (b) is true then since $v_i \in V = \text{null } T + \text{range } T$ so $v_i - w_i \in \text{range } T$ for some $w_i \in \text{null } T = \text{span } (u_1, \dots, u_n)$. Note that $v_1 - w_1, \dots, v_m - w_m$ is linearly dependent as $v_1, \dots, v_m$ is linearly dependent and since $\dim \text{range } T = m$ so $v_1 - w_1, \dots, v_m - w_m$ is basis of range T. Hence if there is $v \in (\text{null } T) \cap (\text{range } T)$ then $v = \sum_{i=1}^{n} \alpha_i u_i = \sum_{j=1}^{m} \beta_j (v_j - w_j)$ which leads to $\beta_j = 0$ as $v_1, \dots, v_m, u_1, \dots, u_n$ is linearly dependent. Thus, $v = 0$ or $(\text{null } T) \cap (\text{range } T) = \{0\}$.

If (c) is true then from theorem 3.2.4 (1C), we find null T+range T is a direct sum. Hence, if $v_1, \ldots, v_m$ is basis of null T, $u_1, \ldots, u_n$ is basis of range T then $v_1, \ldots, v_m, u_1, \ldots, u_n$ is linearly independent. On the other hand, since V is finite-dimensional so dim V = dim range T + dim null $T = m + n$. Thus, $v_1, \ldots, v_m, u_1, \ldots, u_n$ is basis of V, or V = null $T \oplus$ range T.

exer:5C:4 4. Let $V = \mathcal{P}(\mathbf{R})$ and $T \in \mathcal{L}(V)$ so $T(1) = 0, Tx = 0, Tx^i = x^2$ for all $i \geq 2$. Then null T = span $(1, x)$ and range T = span (x^2) so null $T \cap$ range $T = \{0\}$ but $V \neq$ null T + range T.

exer:5C:5 5. If λ is not eigenvalue of T then $T - \lambda I$ is invertible, i.e. range $(T - \lambda I) = V$ for any operator $T \in \mathcal{L}(V)$. Thus, the problem is equivalent to proving that T is diagonalizable iff V = null $(T - \lambda I) \oplus$ range $(T - \lambda I)$ for all eigenvalue λ of T.

Claim 1. If T diagonalizable then V = null $(T - \lambda I) \oplus$ range $(T - \lambda I)$ for every eigenvalue λ of T.

Proof 1. If T is diagonalizable. Let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues of T. For $1 \leq i \leq m$, let $U = \bigoplus_{j \neq i} E(\lambda_j, T)$ and consider $(T - \lambda_i I)$ restricted to U. For any $u \in U, u = \sum_{j \neq i} u_j$ with $u_j \in E(\lambda_j, T)$ then

$$(T - \lambda_i I)u = \sum_{j \neq i} \lambda_j u_j - \lambda_i \sum_{j \neq i} u_j = \sum_{j \neq i} (\lambda_j - \lambda_i) u_j.$$

Thus, from this, we follow $(T - \lambda_i I)|_U \in \mathcal{L}(U)$ and $(T - \lambda_i I)|_U$ is injective, so according to theorem 7.1.2 (5A) then $(T - \lambda_i I)|_U$ is invertible so range $(T - \lambda_i I)|_U = U$. According to exercise 3 (5C) then

$$V = \bigoplus_{i=1}^{m} E(\lambda_i, T) = \text{null } (T - \lambda_i I) \oplus U = \text{null } (T - \lambda_i I) \oplus \text{range } (T - \lambda_i I)|_U.$$

This follows, V = null $(T - \lambda_i I)$ + range $(T - \lambda_i I)$ so again from exercise 3, V = null $(T - \lambda_i I) \oplus$ range $(T - \lambda_i I)$. This is true for all $1 \leq i \leq m$.

Proof 2. Note that if T is diagonalizable then $T - \lambda I$ is also diagonalizable for any $\lambda \in \mathbf{C}$. Hence, according to exercise 1 (5C) then V = null $(T - \lambda I) \oplus$ range $(T - \lambda I)$ for all $\lambda \in \mathbf{C}$, as desired.

Claim 2. If V = null $(T - \lambda I) \oplus$ range $(T - \lambda I)$ for all $\lambda \in \mathbf{C}$ then T is diagonalizable.

Proof. Since V is complex vector space so there exists eigenvalue λ_1 of T. Let U_1 = range $(T - \lambda_1 I)$ be a subspace of V then dim $U_1 <$ dim V since dim null $(T - \lambda_1 I) \geq 1$.

If dim $U_1 \geq 1$ then there exists eigenvalue λ_2 of $T|_{U_1}$. Since V = null $(T - \lambda_1 I) \oplus$ range $(T - \lambda_1 I)$ so according to exercise 3 (5C) then null $(T - \lambda_1 I) \cap$ range $(T - \lambda_1 I) = \{0\}$. Let $v_2 \in U_1$ is eigenvalue of $T|_{U_1}$ corresponding to λ_2. If $\lambda_2 = \lambda_1$ then $v_2 \in$ null $(T - \lambda_1 I)$, a contradiction to the fact that null $(T - \lambda_1 I) \cap$ range $(T - \lambda_1 I) = \{0\}$. Thus, $\lambda_1 \neq \lambda_2$. This follows $E(\lambda_2, T) = E(\lambda_2, T|_{U_1})$.

Let $U_2 = \text{range } (T - \lambda_2 I)|_{U_1}$ then U_2 is subspace of U_1. From previous argument, we know that null $(T-\lambda_2 I) \cap \text{range } (T - \lambda_2 I) = \{0\}$ so null $(T-\lambda_2 I)|_{U_1} \cap \text{range } (T-\lambda_2 I)|_{U_1} = \{0\}$ so $U_1 = \text{null } (T-\lambda_2 I)|_{U_1} \oplus \text{range } (T - \lambda_2 I)|_{U_1}$ according to exercise 3 (5C). Combining all of these, we obtain

$$\begin{aligned} V &= \text{null } (T - \lambda_1 I) \oplus \text{range } (T - \lambda_1 I), \\ &= E(\lambda_1, T) \oplus U_1, \\ &= E(\lambda_1, T) \oplus \text{null } (T - \lambda_2 I)|_{U_1} \oplus \text{range } (T - \lambda_2 I)|_{U_1}, \\ &= E(\lambda_1, T) \oplus E(\lambda_2, T|_{U_1}) \oplus U_2, \\ &= E(\lambda_1, T) \oplus E(\lambda_2, T) \oplus U_2. \end{aligned}$$

We do the same thing for U_2. This keeps going as long as dim $U_i \geq 1$ (since there always exists eigenvalue of an operator on complex vector space according to theorem 7.3.1). At the end, we will obtain $V = \bigoplus_{i=1}^m E(\lambda_i, T)$ for eigenvalues λ_i of T. From theorem 7.5.2 (5C), this follows T is diagonalizable.

exer:5C:6

6. Since T has dim V distinct eigenvalues so T is diagonalizable, hence according to theorem 7.5.2 (5C), V has basis $v_1, \ldots, v_n$ which are eigenvectors of T corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$. Thus, $v_1, \ldots, v_n$ are also eigenvectors of S analogously corresponding to $\beta_1, \ldots, \beta_n$. Hence, $STv_i = S(\lambda_i v_i) = \beta_i \lambda_i v_i = TSv_i$ for all $1 \leq i \leq n$. This follows $ST = TS$ as $v_1, \ldots, v_n$ is basis of V.

exer:5C:7

7. T has diagonal matrix A with respect to basis $v_1, \ldots, v_n$ of V. Let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues of T. Let S_k be number of vectors v_i so $v_i \in E(\lambda_k, T)$, i.e. number of times λ_k appears on the diagonal A. Note that $v_1, \ldots, v_n$ is linearly independent so $S_k \leq \text{dim } E(\lambda_k, T)$. Hence, we have

$$n = \sum_{k=1}^m S_k \leq \sum_{k=1}^m \text{dim } E(\lambda_k, T) = n.$$

The inequality holds when $S_k = \text{dim } E(\lambda_k, T)$, i.e. λ_k appears on the diagonal A precisely dim $E(\lambda_k, T)$ times.

exer:5C:8

8. If both $T - 2I$ and $T - 6I$ are not invertible, then according to theorem 7.1.2 (5A), 2 and 6 are eigevalues of T, i.e. dim $E(2,T) \geq 1$, dim $E(6,T) \geq 1$. However,

$$\text{dim } E(2,T) + \text{dim } (6,T) + \text{dim } T(8,T) > 5 = \text{dim } \mathbf{F}^5.$$

This contradicts to theorem 7.5.1 (5C). Thus, either $T - 2I$ or $T - 6I$ or both are invertible.

exer:5C:9

9. According to exercise 21 (5A), if λ is not eigenvalue of T then $\frac{1}{\lambda}$ is not eigenvalue of T^{-1}. Hence, $E(\lambda, T) = E(\frac{1}{\lambda}, T^{-1}) = \{0\}$.

   If λ is eigenvalue of T then according to exercise 21 (5A), if $v \in E(\lambda, T)$ then $v \in E(\frac{1}{\lambda}, T^{-1})$. Similarly and we obtain $E(\lambda, T) = E(\frac{1}{\lambda}, T^{-1})$.

exer:5C:10 10. We have null $T = \text{null } (T - 0I) = E(0,T)$. Hence, according to theorem 7.5.1, theo:5.38:5C

$$\sum_{i=1}^{m} \text{dim } E(\lambda_i, T) + \text{dim } E(0,T) \le \text{dim } V = \text{dim range } T + \text{dim null } T.$$

This follows $\sum_{i=1}^{m} \text{dim } E(\lambda_i, T) \le \text{dim range } T$.

exer:5C:11 11. Done.

exer:5C:12 12. Let v_1, v_2, v_3 be eigenvectors of R corresponding to $2, 6, 7$ then v_1, v_2, v_3 is basis of $\mathbf{F}^3$. Similarly, let w_1, w_2, w_3 be eigenvectors of T corresponding to $2, 6, 7$ then w_1, w_2, w_3 is basis of V. Let $S \in \mathcal{L}(\mathbf{F}^3)$ so $Sv_i = w_i$ then S is injective so S is invertible. We have $S^{-1}TSv_i = S^{-1}Tw_i = S^{-1}(\alpha_i w_i) = \alpha_i v_i = Rv_i$. Thus, $S^{-1}TS = R$.

exer:5C:13 13. Let $T, R \in \mathcal{L}(\mathbf{F}^4)$ so $Tv_1 = Rv_1 = 2v_1, Tv_2 = Rv_2 = 6v_2, Tv_3 = Rv_3 = 7v_3$ and $Tv_4 = v_1 + 7v_4, Rv_4 = v_2 + 6v_4$ where $v_1, \ldots, v_4$ is basis of $\mathbf{F}^4$. It's not hard to check that T, R only have $2, 6, 7$ as eigenvalues. We show that for any if there is $S \in \mathcal{L}(\mathbf{F}^4)$ so $SR = TS$ then S is not invertible, which can follow that there doesn't exists invertible operator S so $R = S^{-1}TS$.

Indeed, if $Sv_1 = \sum_{i=1}^{4} \alpha_i v_i$ then

$$TSv_1 = T\left(\sum_{i=1}^{4} \alpha_i v_i\right) = 2\alpha_1 v_1 + 6\alpha_2 v_2 + 7\alpha_3 v_3 + \alpha_4(v_1 + 7v_4).$$

And

$$SRv_1 = 2Sv_1 = 2\sum_{i=1}^{4} \alpha_i v_i.$$

With this, we follow $\alpha_4 = 0$, which leads to $\alpha_2 = \alpha_3 = 0$. Thus, $Sv_1 = \alpha_1 v_1$. Similarly, $Sv_2 = \alpha_2 v_2$ (note that we can't imply $Sv_3 = \alpha_3 v_3$, because of the way we choose $Tv_4 = v_1 + 7v_4$).

Now, consider TSv_4 and SRv_4. If $Sv_4 = \sum_{i=1}^{4} \beta_i v_i$ then

$$TSv_4 = T\left(\sum_{i=1}^{4} \beta_i v_i\right) = 2\beta_1 v_1 + 6\beta_2 v_2 + 7\beta_3 v_3 + \beta_4(v_1 + 7v_4).$$

And

$$SRv_4 = S(6v_4 + v_2) = \alpha_2 v_2 + 6\sum_{i=1}^{4} \beta_i v_i.$$

If $TSv_4 = SRv_4$ then we must have $\beta_4 = \beta_3 = 0$. This follows $Sv_4 = \beta_1 v_1 + \beta_2 v_2$, i.e. $Sv_4 \in \text{span } (Sv_1, Sv_2)$ since $Sv_1 = \alpha_1 v_1, Sv_2 = \alpha_2 v_2$. Hence, S is not invertible.

In conclusion, with such T, R then there does not exists invertible operator S so $R = S^{-1}TS$.

exer:5C:14 14. If T does not have a diagonal matrix for any basis of $\mathbf{C}^3$ then that means $\dim E(6,T) = \dim E(7,T) = 1$ and T only has eigenvalues $6, 7$. This makes the construction for T much easier: For arbitrary basis of $\mathbf{C}^3$ then choose $T \in \mathcal{L}(\mathbf{C}^3)$ so $Tv_1 = 6v_1, Tv_2 = 7v_2, Tv_3 = v_1 + v_2 + 6v_3$.

Indeed, we show that T has only two eigenvalues 6 and 7. If there exists $v = a_1v_1 + a_2v_2 + a_3v_3 \neq 0$ so

$$Tv = \lambda(a_1v_1 + a_2v_2 + a_3v_3) = 6a_1v_1 + 7a_2v_2 + a_3(v_1 + v_2 + 6v_3)$$

Then $\lambda a_3 = 6a_3, \lambda a_1 = 6a_1 + a_3, \lambda a_2 = 7a_2 + a_3$. Hence, if $\lambda \neq 6, 7$ then $a_3 = a_2 = a_1 = 0$, a contradiction.

Next, we show that $\dim E(6,T) = \dim E(7,T) = 1$. Indeed, if there is $v = a_1v_1 + a_2v_2 + a_3v_3$ so $Tv = 6v = 6a_1v_1 + 7a_2v_2 + a_3(v_1 + v_2 + 6v_3)$. This follows, $a_3 = 0, a_2 = 0$. Thus, span $(v_1) = E(6,T)$. Similarly, span $(v_2) = E(7,T)$.

These two claims combining with theorem 7.5.2 (5C) deduce that T is not diagonalizable.

exer:5C:15 15. As claimed in exercise 14 (5C), since T is not diagonalizable, T has only $6, 7$ as eigenvalues. Hence, as 8 is not eigenvalue of T, $T - 8I$ is invertible according to theorem 7.1.2 (5A), which means there exists $(x, y, z) \in \mathbf{C}^3$ so $(T - 8I)(x, y, z) = (17, \sqrt{5}, 2\pi)$.

exer:5C:16 16. (a) Induction on n, if $T^{n-1}(0,1) = (F_{n-1}, F_n)$ then

$$T^n(0,1) = T(F_{n-1}, F_n) = (F_n, F_{n-1} + F_n) = (F_n, F_{n+1}).$$

(b) If there is $(x, y) \neq (0, 0)$ so $T(x, y) = \lambda(x, y) = (y, x + y)$ then $\lambda x = y, \lambda y = x + y$. This follows $y(\lambda^2 - \lambda - 1) = 0$. Hence, eigenvalues of T is $\lambda = \frac{1\pm\sqrt{5}}{2}$.

(c) We have $\left(1, \frac{1-\sqrt{5}}{2}\right) \in E\left(\frac{1-\sqrt{5}}{2}, T\right)$ and $\left(1, \frac{1+\sqrt{5}}{2}\right) \in E\left(\frac{1+\sqrt{5}}{2}, T\right)$.

(d) We have $(0,1) = \frac{1}{\sqrt{5}}\left(\left(1, \frac{1+\sqrt{5}}{2}\right) - \left(1, \frac{1-\sqrt{5}}{2}\right)\right)$. Hence,

$$\begin{aligned} T^n(0,1) &= \frac{1}{\sqrt{5}}\left(T^n\left(1, \frac{1+\sqrt{5}}{2}\right) - T^n\left(1, \frac{1-\sqrt{5}}{2}\right)\right), \\ &= \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n\left(1, \frac{1+\sqrt{5}}{2}\right) - \left(\frac{1-\sqrt{5}}{2}\right)^n\left(1, \frac{1-\sqrt{5}}{2}\right)\right] \end{aligned}$$

This follows $F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$.

(e) It is equivalent to prove that $\left|\frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n\right| < \frac{1}{2}$, which is true since $\left(\frac{\sqrt{5}-1}{2}\right)^n < 1 < \frac{\sqrt{5}}{2}$.

8. Chapter 6: Inner Product Spaces

8.1. 6A: Inner Products and Norms

theo:6.14:6A **Theorem 8.1.1** (6.14) Suppose $u, v \in V$, with $v \neq 0$. Set $c = \dfrac{\langle u, v\rangle}{\|v\|^2}$ and $w = u - \dfrac{\langle u, v\rangle}{\|v\|^2}v$ then $\langle w, v\rangle = 0$ and $u = cv + w$.

theo:6.15:6A **Theorem 8.1.2** (6.15, Cauchy-Schwarz Inequality) Suppose $u, v \in V$. Then $|\langle u, v\rangle| \leq \|u\| \|v\|$. The inqualitiy holds when one of u, v is a scalar multiple of the other.

8.2. Exercises 6A

exer:6A:1 1. Definiteness is not true, i.e. $|1 \cdot 0| + |1 \cdot 0| = 0$ but $(1, 0) \neq (0, 0)$.

exer:6A:2 2. For $(x_1, y_1, z_1) \in \mathbf{R}^3$ so $x_1 y_1 < 0$ then $\langle (x_1, y_1, z_1), (x_1, y_1, z_1)\rangle = 2x_1 y_1 < 0$, a contradiction to condition of positivity in definition of inner product.

exer:6A:3 3. (INCORRECT - NEED TO CHECK) According to definition of inner product then $\langle v, v\rangle > 0$ for $v \neq 0$, hence any inner product on V satisfying the old definition will also satisfies the new definition.

   Conversely, consider an inner product on V that satisfies the new definition, there exists $v \in V, v \neq 0$ so $\langle v, v\rangle > 0$. Note that with new definition, we can't define norm of vector, however, we can avoid this notation by only woking on $\langle u, u\rangle$ rather than $\sqrt{\langle u, u\rangle}$. With this in mind, we can check that the orthogonal decomposition strill works in the new definition, but only for v, as shown:

   For any $u \in V$ then with $c = \dfrac{\langle u, v\rangle}{\langle v, v\rangle}$ and $w = u - \dfrac{\langle u, v\rangle}{\langle v, v\rangle}v$ then $\langle w, v\rangle = 0$ and $u = cv + w$.

   Hence, with this, we have $\langle u, u\rangle = \langle cv + w, cv + w\rangle = c^2 \langle v, v\rangle + \langle w, w\rangle$

exer:6A:4 4. (a) $\langle u + v, u - v\rangle = \|u\|^2 - \langle u, v\rangle + \langle v, u\rangle - \|v\|^2$. Since we're talking about real inner product, so $\langle u, v\rangle = \langle v, u\rangle$, done.

   (b) This is deduced from (a).

   (c) A rhombus made by two vectors $u, v \in \mathbf{R}^2$ with same norm has two diagonals $u-v, u+v$. From (b), we find $u+v$ is orthogonal to $u-v$, which follows two diagonals of the rhombus are perpendicular to each other.

exer:6A:5 5. If there exists $v \in V, v \neq 0$ so $Tv = \sqrt{2}v$ then $\|Tv\|^2 = \langle \sqrt{2}v, \sqrt{2}v\rangle = 2\|v\|^2 > \|v\|^2$, a contradiction. Thus there does not exists $v \in V, v \neq 0$ so $Tv = \sqrt{2}v$, i.e. $\sqrt{2}$ is not eigenvalue of T, which follows that $T - \sqrt{2}I$ is invertible from theorem 7.1.2 theo:5.6:5A under assumption that V is finite-dimensional.

exer:6A:6 6. If $\langle u, v\rangle = 0$ then according to Pythagorean's theorem, $\|u + av\|^2 = |a|^2 \|v\|^2 + \|u\|^2 \geq \|u\|^2$ for all $a \in \mathbf{F}$.

If $\|u + av\|^2 \geq \|u\|^2$ for all $a \in \mathbf{F}$ then it is equivalent to $0 \leq a^2 \|v\|^2 + 2a\mathrm{Re}\,\langle v, u\rangle$. If $\langle v, u\rangle \neq 0$ then we can choose a so $0 > a^2 \|v\|^2 + 2a\mathrm{Re}\,\langle v, u\rangle$, a contradiction. Thus, $\langle v, u\rangle = 0$.

exer:6A:7 7. Since $a, b \in \mathbf{R}$ so $\|au + bv\|^2 = |a|^2 \|u\|^2 + 2ab\,\langle u, v\rangle + |b|^2 \|v\|^2$, similarly, $\|bu + av\|^2 = |b|^2 \|u\|^2 + 2ab\,\langle u, v\rangle + |a|^2 \|v\|^2$. Hence, $\|au + bv\|^2 = \|bu + av\|^2$ iff $|a|^2 \|u\|^2 + |b|^2 \|v\|^2 = |b|^2 \|u\|^2 + |a|^2 \|v\|^2$ iff $\left(|a|^2 - |b|^2\right)\left(\|u\|^2 - \|v\|^2\right) = 0$ for all $a, b \in \mathbf{R}$ iff $\|v\| = \|u\|$.

exer:6A:8 8. We have $|\langle u, v\rangle| = 1 = \|u\|\,\|v\|$ so according to theorem 8.1.2, $u = av$ for $a \in \mathbf{F}$. From here we find $a = 1$ or $u = v$.

exer:6A:9 9. Let $\|u\| = a, \|v\| = b$ then we have according to theorem 8.1.2 (6A), we have $|\langle u, v\rangle| \leq ab$ so

$$(1 - |\langle u, v\rangle|)^2 \geq (1 - ab)^2 \geq (1 - a^2)(1 - b^2).$$

exer:6A:10 10. Using theorem 8.1.1 (6A), we have with $c = \dfrac{\langle (1,2), (1,3)\rangle}{\|(1,3)\|^2} = \dfrac{7}{10}$ then $u = \dfrac{7}{10}(1,3)$ and $v = (1,2) - \dfrac{7}{10}(1,3) = \left(\dfrac{3}{10}, \dfrac{-1}{10}\right)$.

exer:6A:11 11. For $x = (a^{1/2}, b^{1/2}, c^{1/2}, d^{1/2}), y = (a^{-1/2}, b^{-1/2}, c^{-1/2}, d^{-1/2}) \in \mathbf{R}^4$ then according to theorem 8.1.2 (6A) we have

$$16 = |\langle x, y\rangle|^2 \leq \|x\|^2 \|y\|^2 = (a + b + c + d)\left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}\right).$$

exer:6A:12 12. Apply theorem 8.1.2 to $(x_1, \ldots, x_n), (1, \ldots, 1) \in \mathbf{R}^n$ with respect to Euclidean inner product.

exer:6A:13 13. Draw triangle formed by u, v and $u - v$. Let θ be angle between u and v then applying law of cosines, we have

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\cos\theta\,\|u\|\,\|v\|.$$

On the other hand, for $u, v \in \mathbf{R}^2$ then

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\,\langle u, v\rangle.$$

Thus, $\langle u, v\rangle = \|u\|\,\|v\|\cos\theta$.

exer:6A:14 14. Since $-1 \leq \cos\theta \leq 1$ so we must have $-1 \leq \dfrac{\langle x, y\rangle}{\|x\|\,\|y\|} \leq 1$, which is true according to theorem 8.1.2.

exer:6A:15 15. Consider $x = (a_1, 2^{1/2}a_2, \ldots, n^{1/2}a_n), y = (b_1, 2^{-1/2}b_2, \ldots, n^{-1/2}b_n) \in \mathbf{R}^2$ then according to theorem 8.1.2 with respect to Euclidean inner product, we have the desired inequality.

exer:6A:16 16. $2\|v\|^2 = \|u+v\|^2 + \|u-v\|^2 - 2\|u\|^2 = 34$ so $\|v\| = \sqrt{17}$.

exer:6A:17 17. We have $\|(x,y)\| = \max\{|x|,|y|\} = \frac{1}{2}(|x+y|+|y-x|)$ so $\|(x,y)\|^2 = \frac{1}{2}(x^2+y^2+|y^2-x^2|)$. If such inner product on $\mathbf{R}^2$ exists, then for any $u,v \in \mathbf{R}^2$ then $\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2+\|v\|^2)$. Let $u=(x,y), v=(a,b)$ then

$$\|u+v\|^2+\|u-v\|^2 = x^2+y^2+a^2+b^2+1/2\left|(x+a)^2-(y+b)^2\right|+1/2\left|(x-a)^2+(y-b)^2\right|,$$

and

$$2(\|u\|^2+\|v\|^2) = x^2+y^2+a^2+b^2+|x^2-y^2|+|a^2-b^2|.$$

Then we must have

$$\left|(x+a)^2-(y+b)^2\right|+\left|(x-a)^2+(y-b)^2\right| = 2|x^2-y^2|+2|a^2-b^2|.$$

However, if $x=6, y=-5, a=3, b=2$ then $LHS = 9^2-3^2+7^2-3^2 = 112$ but $RHS = 2(6^2-5^2+3^2-2^2) = 32$, a contradiction. Thus, such inner product on $\mathbf{R}^2$ does not exist.

exer:6A:18 18. Exercise 19 (next exercise) makes it easier, in which we have

$$\langle (1,1),(1,-1)\rangle = \frac{1}{4}\left(\|(2,0)\|^2 - \|(0,2\|^2\right) = 0.$$

So $(1,1)$ is othorgonal to $(1,-1)$, hence according to Pythagorean's theorem then

$$\|(1,1)+(1,-1)\|^2 = \|(2,0)\|^2 = \|(1,1)\|^2 + \|(1,-1)\|^2.$$

We have $LHS = 2^2 = 4$ and $RHS = 2\|(1,1)\|^2 = 2\cdot 2^{2/p}$. This leads to $p=2$.

If $p=2$ then according to exercise 19, it suffices to show

$$\langle (x,y),(a,b)\rangle = \frac{1}{4}\left[(x+a)^2+(y+b)^2-(x-a)^2-(y-b)^2\right] = xa+yb.$$

is an inner product on $\mathbf{R}^2$. This is obviously true as this inner product is Euclidean inner product.

exer:6A:19 19. We have

$$\begin{aligned}\|u+v\|^2 - \|u-v\|^2 &= \langle u+v,u+v\rangle - \langle u-v,u-v\rangle,\\ &= \left(\|u\|^2+\langle u,v\rangle+\langle v,u\rangle+\|v\|^2\right) - \left(\|u\|^2-\langle u,v\rangle-\langle v,u\rangle+\|v\|^2\right),\\ &= 2\langle u,v\rangle + 2\langle v,u\rangle.\end{aligned}$$

In real inner product, we find $\langle u,v\rangle = \langle v,u\rangle$, which will obtain the desired identity.

exer:6A:20 20. From exercise 19, we know that $\|u+v\|^2 - \|u-v\|^2 = 2\langle u,v\rangle + 2\langle v,u\rangle$. Next, we have

$$\begin{aligned}
\|u+iv\|^2 - \|u-iv\|^2 &= \langle u+iv, u+iv\rangle^2 - \langle u-iv, u-iv\rangle\,,\\
&= \left(\|u\|^2 + \langle u, iv\rangle + \langle iv, u\rangle + \|iv\|^2\right)\\
&\quad - \left(\|u\|^2 + \|iv\|^2 - \langle u, iv\rangle - \langle iv, u\rangle\right),\\
&=2\left[\langle u, iv\rangle + \langle iv, u\rangle\right],\\
&=2i\left[\langle v, u\rangle - \langle u, v\rangle\right].
\end{aligned}$$

Thus,

$$\langle u,v\rangle = \frac{1}{4}\left(\|u+v\|^2 - \|u-v\|^2 + \|u+iv\|^2 i - \|u-iv\|^2 i\right).$$

exer:6A:21 21. We will only prove for the case $\mathbf{F}=\mathbf{R}$, the other case $\mathbf{F}=\mathbf{C}$ is similar (which proof for this case can be seen here) . Define a function $\langle\,,\rangle : U\times U \to \mathbf{R}$ so $\langle u,v\rangle = \frac{1}{4}\left(\|u+v\|^2 - \|u-v\|^2\right)$. Then it's obvious that $\|u\| = \langle u,u\rangle^{1/2}$. It suffices to show that $\langle ;,\rangle$ is an inner product.

Indeed, the positivity and definiteness conditions satisfies since $\langle u,u\rangle = \|u\|^2$ is 0 iff $\|u\|=0$ iff $u=0$. Now we check the condition of additivity in first slot, we have

$$\langle u,v\rangle + \langle w,v\rangle = \frac{1}{4}\left(\|u+v\|^2 + \|w+v\|^2 - \|u-v\|^2 - \|w-v\|^2\right).$$

Since the norm satisfies the parallelogram equality so

$$\begin{aligned}
\|u+v\|^2 + \|w+v\|^2 &= \frac{1}{2}\left(\|w+u+2v\|^2 + \|u-w\|\right)^2,\\
&= \frac{1}{2}\left[\|u-w\|^2 + 2\left(\|w+u+v\|^2 + \|v\|^2\right) - \|w+u\|^2\right],\\
&= \|w+u+v\|^2 + \frac{1}{2}\|u-w\|^2 + \|v\|^2 - \frac{1}{2}\|w+u\|^2
\end{aligned}$$

Similarly, we obtain

$$\|u-v\|^2 + \|w-v\|^2 = \|w+u-v\|^2 + \|v\|^2 + \frac{1}{2}\|u-w\|^2 - \frac{1}{2}\|w+u\|^2.$$

Thus,

$$\langle u,v\rangle + \langle w,v\rangle = \frac{1}{4}\left(\|w+u+v\|^2 - \|w+u-v\|^2\right) = \langle w+u, v\rangle.$$

Next we check the condition of homogeneity in first slot. From the condition of additivity of first slot, we follow that $\langle nu, v\rangle = n\langle u,v\rangle$ for any $n\in\mathbf{Z}$. For $\lambda = a/b \in \mathbf{Q}, a,b\in\mathbf{Z}, b\neq 0$ then for any $u,v\in U$, we have $u/b\in U$ so $b\langle u/b, v\rangle = \langle u,v\rangle$. Hence,

$$\lambda\langle u,v\rangle = \frac{a}{b}\langle u,v\rangle = \left\langle \frac{a}{b}u, v\right\rangle = \langle \lambda u, v\rangle.$$

So far, homogeneity in first slot is true on $\mathbf{Q}$. Since $\|u\|+\|v\| \geq \|u+v\|$ so $(\|u\|+\|v\|)^2 \geq \|u+v\|^2$ so $\|u\|\|v\| \geq \langle u,v\rangle$. From $\|u\|+\|v\|\geq\|u+v\|$ we also follow that the norm is continuous.

Next, consider $\lambda \in C$ and consider sequence $(r_n)_{n\geq 1} \to \lambda$ where $r_i \in \mathbf{Q}$ then we have

$$\langle r_n u, v\rangle - \langle \lambda u, v\rangle = \langle (r_n - \lambda)u, v\rangle \leq \|(r_n - \lambda)u\|\|v\| = |r_n - \lambda|\|u\|\|v\|.$$

Since $\lim_{n\to\infty} |r_n - \lambda|\|u\|\|v\| = 0$ so $\lim_{n\to\infty} \langle r_n u, v\rangle = \langle \lambda u, v\rangle$. Since $r_n \in \mathbf{Q}$ so $\lim_{n\to\infty} \langle r_n u, v\rangle = \langle u, v\rangle \lim_{n\to\infty} r_n = \lambda \langle u, v\rangle$. Thus, $\langle \lambda u, v\rangle = \lambda \langle u, v\rangle$ for all $\lambda \in \mathbf{R}$.

Thus, the function $\langle , \rangle$ is indeed an inner product.

exer:6A:22 22. We need to show that for any $a_1, \ldots, a_n \in \mathbf{R}$ then $(a_1 + \ldots + a_n)^2 \leq n\left(a_1^2 + \ldots + a_n^2\right)$. Consider $x = (a_1, \ldots, a_n), y = (1, 1, \ldots, 1) \in \mathbf{R}^n$ then according to theorem 8.1.2, we have

$$|a_1 + \ldots + a_n| = |\langle x, y\rangle| \leq \|x\|\|y\| = \sqrt{n(a_1^2 + \ldots + a_n^2)}.$$

exer:6A:23 23. We have $\langle (v_1, \ldots, v_m), (v_1, \ldots, v_m)\rangle = \langle v_1, v_1\rangle + \ldots + \langle v_m, u_m, \rangle$ equals 0 iff $\langle v_i, v_i\rangle = 0$ iff $v_i = 0$ for all $1 \leq i \leq m$. One can easily check other conditions in the definition of inner space.

exer:6A:24 24. We have $\langle u, u\rangle_1 = 0$ iff $\langle Su, Su\rangle = 0$ iff $Su = 0$ iff $u = 0$ since S is injective. We also have

$$\langle u + w, v\rangle_1 = \langle S(u + w), Sv\rangle = \langle Su, Sv\rangle + \langle Sw, Sv\rangle = \langle u, v\rangle_1 + \langle w, v\rangle_1.$$

Similarly, $\langle \lambda u, v\rangle_1 = \lambda \langle u, v\rangle_1$ and $\overline{\langle u, v\rangle_1} = \overline{\langle Su, Sv\rangle} = \langle Sv, Su\rangle = \langle v, u\rangle_1$.

exer:6A:25 25. Since S is not injective so there exists $u \in V, u \neq 0$ so $Sv = 0$. Hence, $\langle v, v\rangle_1 = 0$ but $v \neq 0$, which fails the definiteness condition of inner product.

exer:6A:26 26. (a) This is talking about Euclidean inner product space. So

$$\begin{aligned}
\langle f(t), g(t)\rangle' &= \langle (f_1(t), \ldots, f_n(t)), (g_1(t), \ldots, g_n(t))\rangle', \\
&= [f_1(t)g_1(t) + \ldots + f_n(t)g_n(t)]', \\
&= \sum_{i=1}^n \left[f_i'(t)g_i(t) + f_i(t)g_i'(t)\right], \\
&= \sum_{i=1}^n f_i'(t)g_i(t) + \sum_{i=1}^n f_i(t)g_i'(t), \\
&= \left\langle f'(t), g(t)\right\rangle + \left\langle f(t), g'(t)\right\rangle.
\end{aligned}$$

(b) We have $\langle f(t), f(t)\rangle = \|f(t)\|^2 = c^2$ so $\langle f(t), f(t)\rangle' = 0$. On the other hand, from (a) then $\langle f(t), f(t)\rangle' = 2\langle f'(t), f(t)\rangle$ so done.

(c) (DON'T KNOW WITH CURRENT KNOWLEDGE)

exer:6A:27 27. Since $\|w - u\| + \|w - v\|^2 = \frac{1}{2}\left(\|2w - u - v\|^2 + \|u - v\|^2\right)$ so we are done.

exer:6A:28 28. If there is two points $u, v \in C, u \neq v$ closest to $w \in C$ then $\|w - u\| = \|w - v\| \leq \|w - r\|$ for all $r \in C$. Since $\frac{1}{2}(u + v \in C$ so we have $\|w - u\|^2 \leq \left|w - \frac{1}{2}(u + v)\right|^2 = \|w - u\|^2 - \frac{1}{4}\|u - v\|^2$, which follows $\|u - v\|^2 = 0$ or $u = v$, a contradiction. Thus, there is at most one point in C that is closest to w.

exer:6A:29 29. (a) $d(u,v)\geq 0$ and $d(u,v)=0$ iff $\|u-v\|=0$ iff $u=v$; $d(u,v)=d(v,u)$; $d(u,v)+d(v,w)=\|u-v\|+\|v-w\|\geq\|u-w\|=d(u,w)$ according to definition of norm on exercise 21. Thus d is metric on V.

(b) (I couldn't solve this so I did some research and found the proof, note that norm is defined in exercise 21 (6A)). Let $v_1,\dots,v_n$ be basis of V. For all $x\in V, x=\sum_{i=1}^n \alpha_i v_i$, define a norm $\|x\|_1=\sum_{i=1}^n|\alpha_i|$ (can verify that it is indeed a norm). Consider a Cauchy sequence $\{x_k\}_{k\geq 1}$ in the metric space (V,d) with $x_k=\sum_{i=1}^n \alpha_{i,k}v_i$, i.e. for any $\varepsilon>0$ there exists $N>0$ so for all $m,n>N$ we have

$$d(x_m,x_n)=\|x_m-x_n\|=\left\|\sum_{i=1}^n(\alpha_{i,m}-\alpha_{i,n})v_i\right\|<\varepsilon.$$

Since any two norms on some finite-dimensional vector space V over $\mathbf{F}$ are equivalent, there exists pairs of real numbers $0<C_1\leq C_2$ so for all $x\in V$ then $C_1\|x\|_1\leq\|x\|\leq C_2\|x\|_1$. Hence, for any $j=1,\dots,n$, we have

$$\left\|\sum_{i=1}^k(\alpha_{i,m}-\alpha_{i,n})v_i\right\|\geq C_1\sum_{i=1}^n|\alpha_{i,m}-\alpha_{i,n}|\geq C_1|\alpha_{j,m}-\alpha_{j,n}|.$$

This follows that for any $\varepsilon>0$, there exists $N>0$ so for all $m,n>N$ then $|\alpha_{j,m}-\alpha_{j,n}|<\varepsilon$. This follows that $\{\alpha_{j,i}\}_{i\geq 1}$ is a Cauchy sequence in $\mathbf{F}$. Since $\mathbf{F}$ is complete (see page 2 in here for proof of case $\mathbf{F}=\mathbf{R}$) so for any $j=1,\dots,n$, $\{\alpha_{j,i}\}_{i\geq 1}$ converges to β_j. Consider $a=\sum_{i=1}^n\beta_i v_j$. We show that Cauchy sequence $\{x_k\}_{k\geq 1}$ converges, i.e. for any $\varepsilon>0$, there exists $N>0$ so for all $n>N$ then $d(x_n,a)=\|x_n-a\|<\varepsilon$. This is true because

$$\|x_n-a\|=\left\|\sum_{i=1}^n(\alpha_{i,n}-\beta_i)v_i\right\|\leq C_2\sum_{i=1}^n|\alpha_{i,n}-\beta_i|<C_2\sum_{i=1}^n\varepsilon_i.$$

(c) It suffices to prove that any Cauchy sequence $\{x_k\}_{k\geq 1}$ in U converges to x in U. This is true if we choose $u_1,\dots,u_n$ as basis of U and prove similarly to (b), which will gives $u\in\operatorname{span}(u_1,\dots,u_n)=U$.

exer:6A:30 30. (Followed from the hint in the book) Let $p=q+\left(1-\|x\|^2\right)r$ for some polynomial r on $\mathbf{R}^n$. It suffices to choose r so p is a harmonic polynomial on $\mathbf{R}^n$, i.e. $\Delta\left(1-\|x\|^2\right)r=\Delta(-q)$. Let V be set of polynomial r on $\mathbf{R}^n$ so $\deg r\leq\deg q$. Hence V is a finite-dimensional vector space with basis $x_1^{m_1}x_2^{m_1}\cdots x_n^{m_n}$ where $\sum_{i=1}^n m_i\leq\deg q$. Define an operator T on V so for $r\in V$ then $Tr=\Delta\left(1-\|x\|^2\right)r=\Delta\left(1-\sum_{i=1}^n x_i^2\right)r$. Consider polynomial r so $Tr=0$, i.e. $h=\left(1-\sum_{i=1}^n x_i^2\right)r$ is harmonic polynomial. Since harmonic polynomial $h(x)=0$ for every $x\in\mathbf{R}^n$ so $\|x\|=1$, we follow $h(x)=0$. Thus, T is injective, which leads to T being surjective. Hence, for polynomial $\Delta(-q)$, there exists polynomial r so $Tr=\Delta\left(1-\|x\|^2\right)r=\Delta(-q)$. By letting $p=q+\Delta\left(1-\|x\|^2\right)r$, we find $\Delta p=0$, i.e. p is a harmonic polynomial on $\mathbf{R}^n$.

exer:6A:31 31. Use exercise 27 (6A), with $a=w-u, b=w-v, c=u-v$ and $d=w-\frac{1}{2}(u+v)$.

8.3. 6B: Orthonormal Bases

theo:6.30:6B **Theorem 8.3.1** (6.30) Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $v \in V$. Then $v = \langle v, e_1 \rangle e_1 + \dots + \langle v, e_2 \rangle$ and $\|v\|^2 = |\langle v, e_1 \rangle|^2 + \dots + |\langle v, e_n \rangle|^2$.

theo:6.31:6B **Theorem 8.3.2** (6.31, Gram-Schmidt Procedure) Suppose $v_1, \dots, v_m$ is a linearly independent list of vectors in V. Let $e_1 = v_1/\|v_1\|$. For $j = 2, \dots, m$, define e_j inductively by

$$e_j = \frac{v_j - \langle v_j, e_1 \rangle e_1 - \dots - \langle v_j, e_{j-1} \rangle e_{j-1}}{\|v_j - \langle v_j, e_1 \rangle e_1 - \dots - \langle v_j, e_{j-1} \rangle e_{j-1}\|}.$$

Then $\text{span}\,(v_1, \dots, v_j) = \text{span}\,(e_1, \dots, e_j)$ for every $1 \le j \le m$.

Gram-Schmidt Procedure helps us to find an orthonormal basis of a vector space. From this, we follow:

theo:6.34:6B **Theorem 8.3.3** (6.34) Every finite-dimensional inner product space has an orthonormal basis.

theo:6.37:6B **Theorem 8.3.4** (6.37) Suppose $T \in \mathcal{L}(V)$. If T has an upper-triangular matrix with respect to some basis of V, then T has an upper-triangular matrix with respect to some orthonormal basis of V.

theo:6.42:6B **Theorem 8.3.5** (6.42, Riesz Representation Theorem) Suppose V is finite-dimensional and φ is a linear functional on V. Then there is a unique vector $u \in V$ such that $\varphi v = \langle v, u \rangle$. In particular, $u = \overline{\varphi(e_1)}e_1 + \dots + \overline{\varphi(e_n)}e_n$ for ANY orthonormal basis $e_1, \dots, e_n$ of V.

8.4. Exercises 6B

exer:6B:1

1. (a) We have $\langle (\cos\theta, \sin\theta), (-\sin\theta, \cos\theta) \rangle = 0$ and $\|(\cos\theta, \sin\theta)\| = \|(-\sin\theta, \cos\theta)\| = \sqrt{(\cos\theta)^2 + (\sin\theta)^2} = 1$. Similarly with the other orthonormal basis.

   (b) If $(x_1, y_1), (x_2, y_2)$ is an orthonormal basis of $\mathbf{R}^2$ then $x_1x_2 + y_1y_2 = 0$ and $x_1^2 + y_1^2 = x_2^2 + y_2^2 = 1$. Since $|x_1| \le 1$ so there exists $\theta \in \mathbf{R}$ so $x_1 = \cos\theta$, which follows $y_1^2 = (\sin\theta)^2$. WLOG, say $y_1 = \sin\theta$ and $x_1 \neq 0$ then $x_2/y_2 = -y_1/x_1 = -\tan\theta$. Hence, $x_2^2 + y_2^2 = y^2\left((\tan\theta)^2 + 1\right) = 1$ so $y_2^2 = (\cos\theta)^2$. WLOG $y_2 = \cos\theta$ then $x_2 = -\sin\theta$. Thus, $(x_1, y_1), (x_2, y_2)$ is $(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)$.

exer:6B:2

2. If $v \in \text{span}\,(e_1, \dots, e_n)$ then it's true according to theorem 8.3.1. theo:6.30:6B

   Conversely, if $\|v\|^2 = |\langle v, e_1 \rangle|^2 + \dots + |\langle v, e_n \rangle|^2$. Let $u = \overline{\langle v, e_1 \rangle}e_1 + \dots + \overline{\langle v, e_n \rangle}e_n$ then we have

   $$\langle u, v \rangle = \langle v, u \rangle = |\langle v, e_1 \rangle|^2 + \dots + |\langle v, e_n \rangle|^2 = \|v\|^2.$$

   Hence, $\langle v, u \rangle = \|v\|^2$. On the other hand, we also have $\|u\|^2 = \|v\|^2$. This follows $\|u - v\|^2 = \|u\|^2 + \|v\|^2 - \langle u, v \rangle - \langle v, u \rangle = 0$, hence $v = u \in \text{span}\,(e_1, \dots, e_n)$.

exer:6B:3 3. According to the proof of theorem 8.3.4 (6B), it suffices to apply the Gram-Schmidt Procedure 8.3.2 to $(1,0,0),(1,1,1),(1,1,2)$. We have $e_1=(1,0,0)$ and

$$(1,1,1)-\langle(1,0,0),(1,1,1)\rangle(1,0,0)=(0,1,1)$$

with $\|(0,1,1)\|=\sqrt{2}$ so $e_2=\left(0,\sqrt{\frac{1}{2}},\sqrt{\frac{1}{2}}\right)$. We have

$$\begin{aligned} e_2-\langle(1,1,2),e_1\rangle e_1-\langle(1,1,2),e_2\rangle e_2 &= (1,1,2)-(1,0,0)-\frac{3}{\sqrt{2}}\left(0,\sqrt{\frac{1}{2}},\sqrt{\frac{1}{2}}\right), \\ &= (0,1,2)-(0,3/2,3/2)=(0,-1/2,1/2). \end{aligned}$$

Since $\|(0,-1/2,1/2)\|=\sqrt{\frac{1}{2}}$ so $e_3=\left(0,-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)$.

exer:6B:4 4. For postive integers i,j so $i>j$, we have

$$\begin{aligned} \int_{-\pi}^{\pi}\frac{\cos ix\cos jx}{\pi}dx &= \frac{2}{\pi}\int_{-\pi}^{\pi}[\cos(i-j)x-\cos(i+j)x]\,dx, \\ &= \frac{2}{\pi}\left[\frac{\sin(i-j)x}{i-j}-\frac{\sin(i+j)x}{i+j}\right]_{-\pi}^{\pi}=0. \end{aligned}$$

$$\begin{aligned} \int_{-\pi}^{\pi}\frac{\sin ix\sin jx}{\pi}dx &= \frac{2}{\pi}\int_{-\pi}^{\pi}[\cos(i-j)x-\cos(i+j)x]\,dx, \\ &= \frac{2}{\pi}\left[\frac{\sin(i-j)x}{i-j}-\frac{\sin(i+j)x}{i+j}\right]_{-\pi}^{\pi}=0. \end{aligned}$$

$$\begin{aligned} \int_{-\pi}^{\pi}\frac{\sin ix\cos jx}{\pi} &= \frac{2}{\pi}\int_{-\pi}^{\pi}[\sin(i+j)x+\sin(i-j)x]\,dx, \\ &= \frac{2}{\pi}\left[\frac{\cos(i+j)x}{-i-j}+\frac{\cos(i-j)x}{j-i}\right]_{-\pi}^{\pi}=0. \end{aligned}$$

$$\int_{-\pi}^{\pi}\frac{\cos ix}{\sqrt{2\pi}}dx=\frac{1}{\sqrt{2\pi}}\left[\frac{\sin ix}{i}\right]_{-\pi}^{\pi}=0.$$

$$\int_{-\pi}^{\pi}\frac{\sin ix}{\sqrt{2\pi}}dx=\frac{1}{\sqrt{2\pi}}\left[\frac{\cos ix}{-i}\right]_{-\pi}^{\pi}=0.$$

$$\int_{-\pi}^{\pi}\frac{\sin ix\cos ix}{\pi}dx=\frac{1}{2\pi}\int_{-\pi}^{\pi}\sin(2ix)\,dx=0.$$

Similarly, we can find that

$$\left\|\frac{1}{\sqrt{2\pi}}\right\|=\left\|\frac{\cos ix}{\sqrt{\pi}}\right\|=\left\|\frac{\sin ix}{\sqrt{\pi}}\right\|=1.$$

exer:6B:5

5. We have $\|1\|^2 = \int_0^1 1 \; dx = 1$ so $e_1 = 1$. We have
$$x - \langle x, e_1 \rangle e_1 = x - \int_0^1 x \; dx = x - \frac{1}{2}.$$
And
$$\left\| x - \frac{1}{2} \right\|^2 = \int_0^1 \left(x - \frac{1}{2} \right)^2 \; dx = \int_0^1 \left(x^2 - x + \frac{1}{4} \right) \; dx = \frac{1}{3} - \frac{1}{2} + \frac{1}{4} = \frac{1}{12}$$
so $e_2 = (2x-1)\sqrt{3}$. We have
$$\begin{aligned} x^2 - \langle x^2, e_1 \rangle e_1 - \langle x^2, e_2 \rangle e_2 &= x^2 - \int_0^1 x^2 \; dx - 3(2x-1) \int_0^1 x^2(2x-1) \; dx, \\ &= x^2 - \frac{1}{3} - \frac{1}{2}(2x-1), \\ &= x^2 - x - \frac{1}{6}. \end{aligned}$$
Since
$$\begin{aligned} \left\| x^2 - x - \frac{1}{6} \right\|^2 &= \int_0^1 \left(x^2 - x - \frac{1}{6} \right)^2 dx, \\ &= \int_0^1 \left(x^4 - 2x^3 + \frac{2}{3}x^2 + \frac{1}{3}x + \frac{1}{36} \right) dx, \\ &= \frac{7}{60}. \end{aligned}$$
Thus, $e_3 = \frac{2\sqrt{105}}{7} \left(x^2 - x - \frac{1}{6} \right)$.

exer:6B:6

6. Since differentiation operator on $\mathcal{P}_2(\mathbf{R})$ has an upper-triangular matrix with respect to the basis $1, x, x^2$ so according to proof of theorem 8.3.4 (6B), the answer is in exercise 5.

exer:6B:7

7. Define a linear functional T on $\mathcal{P}_2(\mathbf{R})$ so $Tp = p\left(\frac{1}{2}\right)$. Consider the inner product $\langle p, q \rangle = \int_0^1 p(x)q(x) \; dx$ then according to theorem 8.3.5 (6B), if $Tp = \langle p, q \rangle$ then $q(x) = e_1(1/2)e_1 + e_2(1/2)e_2 + e_3(1/2)e_3$ where e_1, e_2, e_3 is orthonormal basis of $\mathcal{P}_2(\mathbf{R})$, which can be found in exercise 5 (6B).

exer:6B:8

8. Similarly, we find
$$\begin{aligned} q(x) = &\left(\int_0^1 (\cos \pi x) dx \right) + \left(\int_0^1 (2x-1) \cos \pi x \; dx \right) 3(2x-1), \\ &+ \left(\int_0^1 \left(x^2 - x - \frac{1}{6} \right) \cos \pi x \; dx \right) \frac{60}{7} \left(x^2 - x - \frac{1}{6} \right). \end{aligned}$$

exer:6B:9

9. If the Gram-Schmidt Procedure is applied to a list of vectors $v_1, \ldots, v_m$ that is not linearly independent. WLOG say that $v_1, \ldots, w_{m-1}$ is linearly independent and $v_m \in \operatorname{span}(v_1, \ldots, v_{m-1})$ then we can't define e_m because $\|v_m - \langle v_m, e_1 \rangle e_1 - \cdots - \langle v_m, e_{m-1} \rangle e_{m-1}\| = 0$.

exer:6B:10 10. Since span (v_1) = span (e_1) so $v_1 = \langle v_1, e_1\rangle e_1$ or $e_1 = \frac{v_1}{\langle v_1, e_1\rangle}$. Observe that $\|e_1\| = 1$ so that means $|\langle v_1, e_1\rangle| = \|v_1\|$, which leads to $\langle v_1, e_1\rangle = \pm\|v_1\|$. Thus, there are two choices for e_1. Similarly, from $v_j = \langle v_j, e_1\rangle e_1 + \cdots + \langle v_j, e_j\rangle e_j$ we find $\langle v_j, e_j\rangle = \pm\|v_j - \langle v_j, e_1\rangle e_1 - \cdots - \langle v_j, e_{j-1}\rangle e_{j-1}\|$ so there are two ways to choose e_j. In total, there are 2^m orthonormal lists $e_1, \ldots, e_m$ of vectors in V.

exer:6B:11 11. Choose a $u \in V, u \neq 0$, let $c = \|u\|_1^2/\|u\|_2^2$. For any $v \in V, v \neq 0$ can be decomposed into $v = ku + w$ with $\langle u, w\rangle_1 = 0 = \langle u, w, \rangle_2$ according to theorem 8.1.1 (6A). Hence,

$$\langle v, u\rangle_1 = k\|u\|_1^2 = ck\|u\|_2^2 = c\langle v, u\rangle_2 .$$

On the other hand, we can also write $u = kv + w$ with $\langle v, w\rangle_1 = 0 = \langle v, w, \rangle_2$. We find $\langle v, u\rangle_1 = \overline{k}\|v\|_1^2$ and $\langle v, u\rangle_2 = \overline{k}\|v\|_2^2$. Combining with the above, we find $\|v\|_1^2 = c\|v\|_2^2$ for any $v \in V$.

Therefore, for any $v, w \in V$, if one of them is 0 then $\langle v, w\rangle_1 = c\langle v, w\rangle_2 = 0$. If $v, w \neq 0$ then $v = kw + u$ with $\langle w, u\rangle_1 = \langle w, u\rangle_2 = 0$. Hence,

$$\langle v, w\rangle_1 = k\|w\|_1^2 = ck\|w\|_2^2 = \langle v, w\rangle_2 .$$

exer:6B:12 12. Since V is finite-dimensional, there exists an othornormal basis $e_1, \ldots, e_n$ of V with respect to inner product $\langle\cdot,\cdot\rangle_2$. We apply Gram-Schmidt Procedure to $e_1, \ldots, e_n$ with respect to inner product $\langle\cdot,\cdot\rangle_1$ and obtain another orthonormal basis $u_1, \ldots, u_n$ of V. With this, we have span $(e_1, \ldots, e_j)$ = span $(u_1, \ldots, u_j)$ for all $1 \leq j \leq n$. According to theorem 8.3.1 (6B) we have $\|v\|_2^2 = \sum_{i=1}^n |\langle v, e_i\rangle|^2$ and

$$\begin{aligned}
\|v\|_1^2 &= \sum_{k=1}^n |\langle v, u_k\rangle|^2 , \\
&= \sum_{k=1}^n \left|\left\langle v, \sum_{i=1}^k \langle u_k, e_i\rangle e_i \right\rangle\right|^2 , \\
&= \sum_{k=1}^n \left|\sum_{i=1}^k \overline{\langle u_k, e_i\rangle} \langle v, e_i\rangle \right|^2 , \\
&\leq \sum_{k=1}^n \left(\sum_{i=1}^k \left|\overline{\langle u_k, e_i\rangle}\right| \cdot |\langle v, e_i\rangle| \right)^2 , \\
&= \sum_{k=1}^n A_k |\langle v, e_k\rangle|^2 + \sum_{1\leq i<j\leq n} B_{i,j} |\langle v, e_i\rangle| \cdot |\langle v, e_j\rangle| .
\end{aligned}$$

In above, $A_k, B_{i,j}$ are constants obtained from $|\langle u_i, e_j\rangle|$ so they will be the same regardless of which $v \in V$ is chosen. Note that $2|\langle v, e_i\rangle \langle v, e_j\rangle| \leq \langle v, e_i\rangle^2 + \langle v, e_j\rangle^2$ so if we choose $C = \sum_{k=1}^n A_k + \sum_{1\leq i<j\leq n} B_{i,j}$ we will obtain $C\|v\|_2^2 \geq \|v\|_1^2$ for all $v \in V$.

exer:6B:13 13. Apply Gram-Schmidt Procedure 8.3.2 (6B) to linearly independent list of vectors $v_1, \ldots, v_m$, we find an orthonormal list $e_1, \ldots, e_m$ so span $(e_1, \ldots, e_j) =$ span $(v_1, \ldots, v_j)$ for all $1 \le j \le m$. Hence, according to theorem 8.3.1 (6B), say $w \in$ span $(v_1, \ldots, v_m)$ then we have $w = \langle w, e_1 \rangle e_1 + \cdots + \langle w, e_m \rangle e_m$ and $v_j = \langle v_j, e_1 \rangle e_1 + \cdots + \langle v_j, e_j \rangle e_j$ so

$$\langle w, v_j \rangle = \sum_{i=1}^{j} \langle w, e_i \rangle \overline{\langle v_j, e_i \rangle} = C_j + \langle w, e_j \rangle \overline{\langle v_j, e_j \rangle}.$$

It suffices to choose $\langle w, e_i \rangle$ so $\langle w, v_j \rangle > 0$ for all $1 \le j \le m$.

For any $1 \le j \le m$, we have $\langle v_j, e_j \rangle \neq 0$, otherwise $v_j \in$ span $(e_1, \ldots, e_{j-1}) =$ span $(v_1, \ldots, v_{j-1})$, a contradiction since $v_1, \ldots, v_m$ is linearly independent. Hence, we can actually choose $\langle w, e_i \rangle$ inductively on i so $\langle w, v_i \rangle > 0$. Thus, there exists w so $\langle w, v_i \rangle > 0$ for all $1 \le i \le m$.

exer:6B:14 14. Since $\|v_i - e_i\|^2 < \frac{1}{n}$ so we follow $\|v_i\|^2 + \frac{n-1}{n} < 2|\langle v_i, e_i \rangle|$ for all $1 \le i \le n$.

Assume the contrary that $v_1, \ldots, v_n$ is linearly dependent, i.e. there exsits not-all-zero $a_1, \ldots, a_n \in \mathbf{F}$ so $\sum_{i=1}^n a_i v_i = 0$. From this, for all $1 \le j \le n$, we obtain $\sum_{i=1}^n a_i \langle v_i, e_j \rangle = \langle 0, e_j \rangle = 0$. Hence, if we consider the linear map $T : \mathbf{F}^{n,1} \to \mathbf{F}^{n,1}$ so $Tx = \mathcal{M}(T)x$ where $\mathcal{M}(T)_{i,j} = \langle v_j, e_i \rangle$ for all $1 \le i, j \le n$. We know that with $x = (a_1, \ldots, a_n)^T, x \neq 0$ then $Tx = 0$ so T is not injective, hence T is not invertible so dim range $T \le n - 1$. Note that according to theorem 5.10.7 (3F), dim range T equals column rank of $\mathcal{M}(T)$ equals row rank of $\mathcal{M}(T)$. Hence, n rows of $\mathcal{M}(T)$ is linearly dependent in $\mathcal{F}^{1,n}$, i.e. there exists not-all-zero $b_1, \ldots, b_n \in \mathbf{F}$ so $\sum_{i=1}^n b_i (\langle v_1, e_i \rangle, \cdots, \langle v_n, e_i \rangle) = 0$. Hence, $\sum_{i=1}^n b_i \langle v_j, e_i \rangle = 0$ for all $1 \le j \le n$ so $|b_j \langle v_j, e_j \rangle| = \left| \sum_{k \neq j} b_k \langle v_j, e_k \rangle \right|$ for all $1 \le j \le n$. On the other hand, we have

$$\begin{aligned}
\left| \sum_{k \neq j} b_k \langle v_j, e_k \rangle \right|^2 &\le \left(\sum_{k \neq j} |b_k| \cdot |\langle v_j, e_k \rangle| \right)^2, \\
&\le \sum_{k \neq j} |b_k|^2 \sum_{k \neq j} |\langle v_j, e_k \rangle|^2, \\
&= \left(\|v_j\|^2 - |\langle v_j, e_j \rangle|^2 \right) \sum_{k \neq j} |b_k|^2, \\
&< \left[\frac{1}{n} - (1 - |\langle v_j, e_j \rangle|)^2 \right] \sum_{k \neq j} |b_k|^2.
\end{aligned}$$

Note that for all $1 \le j \le n$, $\sum_{k \neq j} \|b_k\|^2 \neq 0$, otherwise we will have $b_k = 0$ for all $k \neq j$, which leads to $b_j \langle v_j, e_j \rangle = 0$ or $\langle v_j, e_j \rangle = 0$ since $b_j \neq 0$, a contradiction since $0 < \|v_i\|^2 + \frac{n-1}{n} < 2|\langle v_i, e_i \rangle|$. From this and the above inequality, we obtain

$$\frac{|b_j|^2}{\sum_{k \neq j} |b_k|^2} < \frac{1/n - (1 - |\langle v_j, e_j \rangle|)^2}{|\langle v_j, e_j \rangle|^2}$$

This leads to

$$\frac{|b_j|^2}{\sum_{k=1}^n |b_k|^2} < \frac{1/n - (1 - |\langle v_j, e_j\rangle|)^2}{|\langle v_j, e_j\rangle|^2 + 1/n - (1 - |\langle v_j, e_j\rangle|)^2}.$$

Next, we will show that

$$\frac{1/n - (1 - |\langle v_j, e_j\rangle|)^2}{|\langle v_j, e_j\rangle|^2 + 1/n - (1 - |\langle v_j, e_j\rangle|)^2} \leq \frac{1}{n}.$$

Indeed, the above is equivalent to $(n|\langle v_j, e_j\rangle| - n + 1)^2 \geq 0$, which is true. Therefore, for all $1 \leq j \leq n$ then $\frac{|b_j|^2}{\sum_{k=1}^n |b_k|^2} < \frac{1}{n}$, which will leads to a contradiction if we sums up all the inequalities. Thus, $v_1, \ldots, v_n$ is linearly independent so $v_1, \ldots, v_n$ is basis of V.

exer:6B:15 15. We prove for the general $C_{\mathbf{R}}([a, b])$ and $\varphi(f) = f(c)$ for some $c \in [a, b]$. Assume the contrary, there exists such function g then pick $f(x) = (x - c)^2 g(x)$, we have $f(c) = 0$ and $f(x) \in C_{\mathbf{R}}([a, b])$. Hence, $\int_a^b f(x)g(x)dx = \int_a^b [(x - c)g(x)]^2 \, dx = f(c) = 0$ so this follows $(x - c)g(x) = 0$ or $g(x) = 0$ for $x \in [a, b]$ since $g(x)$ is continuous on $[a, b]$. Hence, $\langle f, g\rangle = 0$ for all $f \in C_{\mathbf{R}}([a, b])$, which is a contradiction if we choose f so $f(c) \neq 0$. Thus, there doesn't exist such function g.

exer:6B:16 16. We induct on $\dim V$. For $\dim V = 1$ then there exists $v \in V$ so $\|v\| = 1$ and $Tv = \lambda v$ with $|\lambda| < 1$. Hence, since $0 < |\lambda| < 1$ so $\lim_{m \to \infty} |\lambda|^m = 0$ so for any $\epsilon > 0$, there exists M so for all $m > M$ then $\|T^m v\| = \|\lambda^m v\| = |\lambda|^m < \epsilon$.

If the problem is true for $\dim V = n - 1$. For $\dim V = n$, since $\mathbf{F} = \mathbf{C}$, T has an upper-triangular matrix with respect to some basis of V according to theorem 7.3.2 (5B), which follows T has an upper-triangular matrix with respect to some orthonormal basis $e_1, \ldots, e_n$ of V according to theorem 8.3.4 (6B). Hence, T is invariant under $U = \operatorname{span}(e_1, \ldots, e_{n-1})$ so according to inductive hypothesis, for $0 < \epsilon < 1$, there exists $m \in \mathbf{Z}, m \geq 1$ so $\|T^m v\| \leq \epsilon \|v\|$ for all $v \in U$.

We can easily show that $T^m e_n = u + \alpha_{n,n}^m$ where $u \in U$. Hence, we can prove inductively that for any positive integer k then $T^{mk} e_n = \sum_{i=0}^{k-2} \alpha_{n,n}^{im} T^{(k-1-i)m} u + \alpha_{n,n}^{m(k-1)} T^m e_n$. Since $\max\{\epsilon, |\alpha_{n,n}|\} < 1$ so $\lim_{x \to \infty} x \max\{\epsilon, |\alpha_{n,n}|\}^x = 0$. Hence, there exists X so for all $x > X$ then $x \max\{\epsilon, |\alpha_{n,n}|\}^x < \frac{\varepsilon}{\|u\|}$ for sufficiently small $\varepsilon > 0$. If we choose large enough k, then for any $0 \leq i \leq m - 2$, we have

$$\begin{aligned}
\|\alpha_{n,n}^{im} T^{(k-1-i)m} u\| &= |\alpha_{n,n}|^{im} \left\| T^{(k-1-i)m} u \right\|, \\
&\leq |\alpha_{n,n}|^{im} \epsilon \left\| T^{(k-2-i)m} u \right\|, \\
&\leq |\alpha_{n,n}|^{im} \epsilon^{k-1-i} \|u\|, \\
&\leq \frac{\varepsilon}{i(m-1) + k - 1}.
\end{aligned}$$

We can simultaneously choose large enough k so $|\alpha_{n,n}|^{m(k-1)}\|T^m e_n\| \le \lambda$ which will lead to

$$\begin{aligned}
\|T^{mk}e_n\| &\le \sum_{i=0}^{k-2} \left\|\alpha_{n,n}^{im}T^{(k-1-i)m}u\right\| + \left\|\alpha_{n,n}^{m(k-1)}T^m e_n\right\|, \\
&\le \frac{(k-1)\varepsilon}{i(m-1)+k-1} + \lambda.
\end{aligned}$$

With this we follows that for any $\varepsilon > 0$, there exists positive integer k so $\|T^{mk}e_n\| \le \varepsilon$. Hence, for any $v \in V, v = \sum_{i=1}^n \beta_i e_i = w + \beta_i e_i$ then

$$\begin{aligned}
\|T^{mk}v\| &\le \|T^{mk}w\| + |\beta_n| \cdot \|T^{mk}e_n\|, \\
&\le \epsilon^{mk}\|w\| + |\beta_n|\varepsilon, \\
&= \epsilon^{mk}\sqrt{|\beta_1|^2 + \cdots + |\beta_{n-1}|^2} + |\beta_n|\varepsilon, \\
&\le \lambda\sqrt{|\beta_1|^2 + \cdots + |\beta_n|^2} = \lambda\|v\|.
\end{aligned}$$

We can choose λ to be really small with proper choice of ϵ and k.

exer:6B:17

17. (a) We show that for $u, v \in V$ then $\Phi u + \Phi v = \Phi(u+v)$. Indeed, we have for any $w \in V$ then

$$(\Phi u + \Phi v)(w) = (\Phi u)(w) + (\Phi v)(w) = \langle w, u\rangle + \langle w, v\rangle = \langle w, u+v\rangle = (\Phi(u+v))(w).$$

Since $\mathbf{F} = \mathbf{R}$ so we also have

$$(\Phi\lambda u)(w) = \langle w, \lambda u\rangle = \lambda\langle w, u\rangle = \lambda(\Phi u)(w).$$

Thus, Φ is a linear map from V to V'.

(b) If $\mathbf{F} = \mathbf{C}$, for $v \in V, V \neq 0$, pick $\lambda = 1 - i$ then $(\Phi\lambda v)(v) = \langle v, \lambda v\rangle = (1+i)\|v\|^2 \neq (1-i)\|v\|^2 = \lambda(\Phi v)(v)$. This Φ is not a linear map.

(c) If $\Phi u = 0$ for some $u \in V$ then $(\Phi u)(u) = \|u\|^2 = 0$ which leads to $u = 0$. Thus, Φ is injective.

(d) Since V is finite-dimensional so according to theorem 5.10.1 (3F) then $\dim V = \dim V'$ so $\dim V' = \dim V = \dim \operatorname{null} \Phi + \dim \operatorname{range} \Phi = \dim \operatorname{range} \Phi$ so Φ is surjective. Thus, Φ is invertible so Φ is an isomorphism from V onto V'.

How is (d) an alternative proof of the Riesz Representation Theorem (8.3.5 (6B)) when $\mathbf{F} = \mathbf{R}$?: Note since Φ is surjective so for any linear functional φ in V', there exists $v \in V$ so $\Phi v = \varphi$, which means $\varphi(u) = (\Phi v)(u) = \langle u, v\rangle$ for any $u \in V$.

8.5. 6C: Orthogonal Complements and Minimization Problems

theo:6.47:6C

Theorem 8.5.1 (6.47) Suppose U is finite-dimensional subspace of V. Then $V = U \oplus U^\perp$.

theo:6.51:6C **Theorem 8.5.2** (6.51) Suppose U is finite-dimensional subspace of V. Then $U = (U^\perp)^\perp$.

def:6.53:6C **Definition 8.5.3** (orthogonal projection)**.** Suppose U is a finite-dimensional subspace of V. The orthogonal projection of V onto U is the operator $P_U \int \mathcal{L}(V)$ defined as follows: For $v \in V$, write $v = u + w$ where $u \in U, w \in U^\perp$. Then $P_U v = u$.

theo:6.56:6C **Theorem 8.5.4** (6.56) Suppose U is a finite-dimensional subspace of V, $v \in V$ and $u \in U$. Then $\|v - P_U v\| \leq \|v - u\|$. The inequality holds iff $u = P_U v$.

theo:6.55:6C **Theorem 8.5.5** (Properties of orthogonal projection P_U) Suppose U is a finite-dimensional subspace of V and $v \in V$. Then

1. $P_U \in \mathcal{L}(V)$.
2. $P_U u = u$ for every $u \in U$ and $P_U w = 0$ for every $w \in U^\perp$.
3. range $P_U = U$ and null $P_U = U^\perp$.
4. $v - P_U v \in U^\perp$.
5. $P_U^2 = P_U$.
6. $\|P_U v\| \leq \|v\|$.
7. For every orthonormal basis $e_1, \ldots, e_m$ of U then $P_U v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_m \rangle e_m$.

8.6. Exercises 6C

exer:6C:1 1. If $v \in \{v_1, \ldots, v_m\}^\perp$ iff $\langle v, v_i \rangle = 0$ for all $1 \leq i \leq m$ iff $\langle v, w \rangle = 0$ for all $w \in$ span $(v_1, \ldots, v_m)$ iff $v \in$ span $(v_1, \ldots, v_m)^\perp$. Thus, $\{v_1, \ldots, v_m\}^\perp =$ span $(v_1, \ldots, v_m)^\perp$.

exer:6C:2 2. If $U = V$ then it's obviously $U^\perp = \{0\}$. Conversely, if $U^\perp = \{0\}$. Let $e_1, \ldots, e_m$ be orthonormal basis of U, if $U \neq V$ then there exists $v \notin U, v \in U$. Apply Gram-Schmidt Procedure 8.3.2 (6B) to linear independent list $e_1, \ldots, e_m, v$ to get a orthonormal list $e_1, \ldots, e_m, e_{m+1}$ which leads to $e_{m+1} \in U^\perp$, a contradiction. Thus, $U = V$.

exer:6C:3 3. According to Gram-Schmidt Procedure then span $(u_1, \ldots, u_m) =$ span $(e_1, \ldots, e_m) = U$ so $e_1, \ldots, e_m$ is orthonormal basis of U. Since each f_i is orthogonal to each e_j so $f_i \in U^\perp$ for each $1 \leq i \leq n$. Hence, $f_1, \ldots, f_n$ is an orthonormal list in $U^\perp$. On the other hand, dim $U^\perp =$ dim $V -$ dim $U = n$ so $f_1, \ldots, f_n$ is orthonormal basis of $U^\perp$.

exer:6C:4 4. According to exercise 3 (6C), apply Gram-Schmidt Procedure to $(1, 2, 3, -4), (-5, 4, 3, 2), (1, 0, 0, 0), (0, 1, 0, 0)$.

exer:6C:5 5. Since V is finite-dimensional so U, U^T is finite-dimensional. Hence, since $U = (U^\perp)^\perp$ according to theorem 8.5.2 (6C), so from theorem 8.5.1 (6C), we have $V = U \oplus U^\perp =$

$(U^{\perp})^{\perp} \oplus U^{\perp}$, i.e. we can represent v uniquely as $v = u + w$ where $u \in U = (U^{\perp})^{\perp}, w \in U^{\perp}$ hence $P_{U^{\perp}}v = w, P_U v = u$ so $(P_{U^{\perp}} + P_U)v = u + w = Iv$. Thus, $P_{U^{\perp}}v = I - P_U$.

exer:6C:6 6. For any $w \in W$ then $P_U P_W w = P_U w = 0$ so $w \in \text{null } P_U = U^{\perp}$, or $\langle w, u \rangle = 0$ for all $u \in U$. Hence, $\langle u, w \rangle = 0$ for all $u \in U$, all $w \in W$.

exer:6C:7 7. Let $U = \text{range } P$ then null $P \subset U^{\perp}$ since every vector in null P is orthogonal to every vector in range P. On the other hand, since $P^2 = P$ so acoording to exercise 4 (5B), we find $V = \text{range } P \oplus \text{null } P$. Hence, each $v \in V$ can be written uniquely as $v = Pv + w$ where $Pv \in U, w \in \text{null } P \subset U^{\perp}$. Thus, since null $P \subset U^{\perp}$ so $P_U v = P_U(Pv + w) = Pv$.

exer:6C:8 8. Let $U = \text{range } P$. We will show that every vector in null P is orthogonal to every vector in range P. Indeed, from $P^2 = P$ and exercise 4 (5B), we find $V = \text{range } P \oplus \text{null } P$. Let $e_1, \dots, e_m$ be orthonormal basis of range P and $f_1, \dots, f_n$ be orthonormal basis of null P then $e_1, \dots, e_m, f_1, \dots, f_n$ is basis of V. It suffices to prove that $\langle e_i, f_j \rangle = 0$ for all $1 \le i \le m, 1 \le j \le n$.

   From $P^2 = P$ we follow $v - Pv \in \text{null } P$ for all $v \in V$, i.e. each v can be represented uniquely as $v = Pv + w$ where $w \in \text{null } P$. Hence, for any $\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_n \in \mathbf{F}$ so $v = \sum_{i=1}^m \alpha_i e_i + \sum_{j=1}^n \beta_j f_j$ then $Pv = \sum_{i=1}^m \alpha_i e_i$. Thus, $\|Pv\|^2 = \sum_{i=1}^m |\alpha_i|^2$ and

$$\|v\|^2 = \|Pv\|^2 + \sum_{j=1}^n |\beta_i|^2 + \sum_{i=1}^m \sum_{j=1}^n 2\text{Re } \alpha_i \overline{\beta_j} \langle e_i, f_j \rangle .$$

   If there exists at least one $\langle e_i, f_j \rangle \neq 0$ then choose $\alpha_x = \beta_y = 0$ for all $x \neq i, y \neq j$ and α_i, β_j so $\alpha_i \overline{\beta_j} = K \overline{\langle e_i, f_j \rangle}$ for some real number $K < 0$ so $2K| \langle e_i, f_j \rangle |^2 + |\beta_j|^2 < 0$. Hence,

$$\|v\|^2 - \|Pv\|^2 = |\beta_j|^2 + 2K| \langle e_i, f_j \rangle |^2 < 0.$$

   This contradicts to assumption that $\|v\| \ge \|Pv\|$. Thus, we must have $\langle e_i, f_j \rangle = 0$ for every $1 \le i \le m, 1 \le j \le n$. Thus, every vector in null P is orthogonal to every vector in range $P = U$, or null $P \subset U^{\perp}$. Combining with $v = Pv + w$ for any $v \in V$, we find $P_U v = P_U(Pv + w) = Pv$ for all $v \in V$.

exer:6C:9 9. Since U is finite-dimensional subspace of V so according to theorem 8.5.1 (6C) we have $V = U \oplus U^{\perp}$, or each v can be written uniquely as $v = u + w$ for $u \in U, w \in U^{\perp}$.

   If U is invariant under T then we have $Tu \in U$ so for any $v \in V, v = u + w$ then $P_U T_U v = P_U Tu = Tu = TP_U v$ which leads to $P_U T P_U = TP_U$. Conversely, if $P_U T P_U u = TP_U u$ or $P_U Tu = Tu$ then that means $Tu \in V$ for all $u \in U$. Thus, U is invariant under T.

exer:6C:10 10. Since V is finite-dimensional so according to theorem 8.5.1 (6C) we have $V = U \oplus U^{\perp}$, i.e. each v can be represented uniquely as $v = u + w$ with $u \in U, w \in U^{\perp}$.

   If U and $U^{\perp}$ are both invariant under T then for any $v \in V, Tv = Tu + Tw$ with $Tu \in U, Tw \in U^{\perp}$. Hence, $P_U T v = P_U(Tu + Tw) = Tu = TP_U v$ so $P_U T = TP_U$. Conversely, if $P_U T = TP_U$ then for any $u \in U$ we have $P_U Tu = TP_U u = Tu$ so $Tu \in U$. This follows U is invariant under T. For any $w \in U^{\perp}$ we have $P_U Tw = TP_U w = T(0) = 0$ so $Tw \in \text{null } P_U = U^{\perp}$. Thus, $U^{\perp}$ is invariant under T.

exer:6C:11 11. Such u is exactly $P_U(1,2,3,4)$ according to theorem 8.5.4 (6C). We do this step by step:

- Apply Gram-Schmidt Procedure to $(1,1,0,0),(1,1,1,2)$ and obtain orthonormal basis of U which is e_1, e_2.
- Calculate $\langle (1,2,3,4), e_1\rangle, \langle (1,2,3,4), e_2\rangle$ (note that since $V=\mathbf{R}^4$ so we're talking about Euclidean inner product).
- $P_U(1,2,3,4) = \langle (1,2,3,4), e_1\rangle e_1 + \langle (1,2,3,4), e_2\rangle e_2$ is the answer.

exer:6C:12 12. Define subspace U of $\mathcal{P}_3(\mathbf{R})$ as $U = \{p(x)\in\mathcal{P}_3(\mathbf{R}) \deg p \leq 3, p'(0)=0, p(0)=0\}$. This follows $U=\{ax^3+bx^2 : a,b\in\mathbf{R}\} = \text{span}\,(x^3,x^2)$. Hence, similarly to previous exercise, we find orthonormal basis of U from basis x^3, x^2 and inner product on $\mathcal{P}_3(\mathbf{R})$:

$$\langle p,q\rangle = \int_{-\pi}^{\pi} p(x)q(x)dx.$$

From that we find $P_U(3x+2) = \langle 3x+2, e_1\rangle e_1 + \langle 3x+2, e_2\rangle e_2$.

exer:6C:13 13. It is exercise 6.58 in the book.

exer:6C:14 14. (a) If $U^\perp \neq \{0\}$, then we show that for any $g\in U^\perp, g\neq 0$, we have $xg(x)\in U^\perp$. Indeed, for any $f\in U$ then $f(x)x\in U$ so

$$\langle f(x), xg(x)\rangle = \int_{-1}^{1} [xf(x)]\, g(x)dx = \langle xf(x), g(x)\rangle = 0.$$

Hence, $xg(x)\in U^\perp$. On the other hand, $(xg(x))(0)=0$ so $xg(x)\in U$. Thus, we must have $xg(x)=0$ or $g(x)=0$, a contradiction. Thus, $U^\perp=\{0\}$. Another proof for (a) can be seen here.

(b) Since $U^\perp=\{0\}$ so theorems 8.5.1, 8.5.2 (6C) do not hold in this case where $C_{\mathbf{R}}([-1,1])$ is infinite-dimensional.

9. Chapter 7: Operators on Inner Product Spaces

9.1. 7A: Self-adjoint and Normal Operators

V, W are assumed to be finite-dimensional inner product spaces over $\mathbf{F}$.

def:7.2:7A **Definition 9.1.1** (adjoint)**.** Suppose $T \in \mathcal{L}(V, W)$. The adjoint of T is the function $T^* : W \to V$ such that $\langle Tv, w\rangle = \langle v, T^*w\rangle$.

def:7.11:7A **Definition 9.1.2** (self-adjoint operator)**.** An operator $T \in \mathcal{L}(V)$ is called **self-adjoint** or **Hermitian** if $T = T^*$.

theo:7.7:7A **Theorem 9.1.3** (7.7) Suppose $T \in \mathcal{L}(V, W)$ then:

(a) null $T^* = (\text{range } T)^\perp$.

(b) range $T^* = (\text{null } T)^\perp$.

(c) null $T = (\text{range } T^*)^\perp$.

(d) range $T = (\text{null } T^*)^\perp$.

From theorem 9.1.3, we find that if T is self-adjoint, then that means null $T = (\text{range } T)^\perp$. If adding the condition $T^2 = T$, we obtain that T is an orthogonal projection 8.5.3.

theo:7.10:7A **Theorem 9.1.4** (7.10) Let $T \in \mathcal{L}(V, W)$. Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $f_1, \dots, f_m$ is an orthonormal basis of W. Then $\mathcal{M}(T^*, (f_1, \dots, f_m), (e_1, \dots, e_n))$ is the conjugate transpose of $\mathcal{M}(T, (e_1, \dots, e_n), (f_1, \dots, f_m))$.

From theorem 9.1.4 (7A), we find that for self-adjoint operator A, i.e. $A = A^*$ then matrix of A with respect to some orthonormal basis is equal to its conjugate transpose. This is called **Hermitian** matrix. If we're talking about real inner product space then matrix of self-adjoint operator A (wrt orthonormal basis) is equal to its own transpose. This kind of matrix is called **real symmetric** matrix.

theo:7.13:7A **Theorem 9.1.5** (7.13) Every eigenvalue of a self-adjoint operator is real.

theo:7.21:7A **Theorem 9.1.6** Suppose $T \in \mathcal{L}(V)$ is normal and $v \in V$ is eigenvector of T with eigenvalue λ. Then v is also eigenvector of T^* with eigenvalue $\overline{\lambda}$.

theo:7.22:7A **Theorem 9.1.7** (7.22) Suppose $T \in \mathcal{L}(V)$ is normal. Then eigenvectors of T corresponding to distinct eigenvalues are orthogonal.

How to find the adjoint T^* of $T \in \mathcal{L}(V, W)$?: There are two ways:

1. Directly: Find T^* from $\langle Tv, w\rangle = \langle v, T^*w\rangle$. See exercise 6 (7D) or exercise 1 (7A) as examples.
2. Indirectly: Find orthonormal bases of V and W. Write $\mathcal{M}(T)$ with respect two these two bases then $\mathcal{M}(T^*)$ is the conjugate transpose of $\mathcal{M}(T)$ according to theorem 9.1.4 (7A).

9.2. Exercises 7A

exer:7A:1 1. For any $(a_1, \ldots, a_n) \in \mathbf{F}^n$ then

$$\begin{aligned}\langle (a_1,\ldots,a_n), T^*(z_1,\ldots,z_n)\rangle &= \langle T(a_1,\ldots,a_n),(z_1,\ldots,z_n)\rangle,\\ &= \langle (0,a_1,\ldots,a_{n-1}),(z_1,\ldots,z_n)\rangle,\\ &= \sum_{i=1}^{n-1} a_i\overline{z_{i+1}},\\ &= \langle (a_1,\ldots,a_n),(z_2,\ldots,z_n,0)\rangle.\end{aligned}$$

Thus, $T^*(z_1,\ldots,z_n) = (z_2,\ldots,z_n,0)$.

exer:7A:2 2. According to theorem 9.1.3 (7A), we have range $(T^* - \overline{\lambda}I) = (\text{null }(T-\lambda I))^\perp$ so that means

$$\begin{aligned}\dim\text{ null }(T^*-\overline{\lambda}I) &= \dim V - \dim\text{ range }(T^*-\overline{\lambda}I),\\ &= \dim V - (\dim\text{ null }(T-\lambda I))^\perp,\\ &= \dim\text{ null }(T-\lambda I).\end{aligned}$$

Thus, λ is eigenvalue of T iff dim null $(T-\lambda I) \geq 1$ iff dim null $(T^* - \overline{\lambda}I) \geq 1$ iff $\overline{\lambda}$ is eigenvalue of T^*.

exer:7A:3 3. U is invariant under T iff $Tu \in U$ for every $u \in U$ iff for every $u \in U$, $\langle u, T^*v\rangle = \langle Tu, v\rangle = 0$ for all $v \in U^\perp$ iff $T^*v \in U^\perp$ for every $v \in U^\perp$ iff $U^\perp$ is invariant under T^*.

exer:7A:4 4. T is injective iff null $T = \{0\}$ iff (range $T^*)^\perp = \{0\}$ (according to theorem 9.1.3) iff range $T^* = \{0\}^\perp = V$ (according to theorem 8.5.2 6C) iff T^* is surjective. Similarly, T is surjective iff T^* is injective.

exer:7A:5 5. We have dim range T = dim V − dim null T = dim (null $T)^\perp$ but according to theorem 9.1.3 then dim (null $T)^\perp$ = dim range T^* so dim range T = dim range T^*. Since dim range T = dim (null $T^*)^\perp$ = dim W−dim null T^* and dim range T^* = dim (null $T)^\perp$ = dim V − dim null T so we obtain the desired equality.

exer:7A:6 6. (a) Write out orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ using Gram-Schmidt Procedure to $1, x, x^2$ then write out $\mathcal{M}(T)$ with respect to this orthonormal basis.

   (b) This is not a contradiction because the matrix of T is created from nonorthonormal basis of $\mathcal{P}_2(\mathbf{R})$.

exer:7A:7

7. Since S, T are self-adjoint so for any $v, w \in V$ then $\langle STv, w\rangle = \langle Tv, Sw\rangle = \langle v, TSw\rangle$. Thus, ST is self-adjoint iff for any $v, w \in V$ then $\langle STv, w\rangle = \langle v, STw\rangle$ iff $\langle v, TSw\rangle = \langle v, STw\rangle$ iff $ST = TS$.

exer:7A:8

8. We have 0 is a self-adjoint operator on V. For any self-adjoint operators S, T on V then $S+T = S^*+T^* = (S+T)^*$ so $S+T$ is self-adjoint. For any $\lambda \in \mathbf{R}$ then $\lambda T = \overline{\lambda}T^* = (\lambda T)^*$ so λT is self-adjoint. Thus, the set of self-adjoint operators on V is a subspace of $\mathcal{L}(V)$.

exer:7A:9

9. For any self-adjoint operator $T \neq 0$ on V, then with $\lambda = 1-i$, we have $\lambda T = (1-i)T = (1-i)T^* \neq (1+i)T^* = (\lambda T)^*$.

exer:7A:10

10. We show that the set of normal operators on V is not closed under addition. It suffices to prove this when dim $V = 2$. Note that $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is the matrix of a normal operator iff $b^2 = c^2$ and $ac + bd = ab + cd$. Hence, with two normal operators S, T so $\mathcal{M}(S) = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$ and $\mathcal{M}(T) = \begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}$ then $\mathcal{M}(S+T) = \begin{pmatrix} 3 & -1 \\ 5 & 5 \end{pmatrix}$ so $S+T$ is not a normal operator. Thus, for dim $V = 2$, set of normal operators on V is not a subspace of $\mathcal{L}(V)$. For dim $V > 2$, we can choose $\mathcal{M}(X) = \begin{pmatrix} \mathcal{M}(S) & \\ & 0 \end{pmatrix}$ (i.e. the 2-by-2 matrix at top left of $\mathcal{M}(X)$ is exactly $\mathcal{M}(S)$ while the remaining entries of $\mathcal{M}(X)$ are 0) and $\mathcal{M}(Y) = \begin{pmatrix} \mathcal{M}(T) & \\ & 0 \end{pmatrix}$.

exer:7A:11

11. If P is self-adjoint then $P = P^*$ so according to theorem 9.1.3 (7A) then null P = null $P^* = (\text{range } T)^\perp$, i.e. every vector in null P is orthogonal to every vector in range P, which brings us back to exercise 7 (6C), which proves the existence of U.

   Conversely, if such subspace U exists then since $P = P_U$ so range P = range $P_U = U$. This follows null P = null $P_U = U^\perp$. Thus, range $P = (\text{null } P)^\perp$ = range P^* and similarly, null P = null P^*. On the other hand, since $P = P^2$ so $P^* = (P^2)^* = (P^*)^2$ so according to exercise 4 (5B) we find every v can be represented uniquely as $v = P^*v + w$ where $w \in$ null P^*. We also have $v = Pv + w$ since $P = P^2$ and null P^* = null P and range P = range P^*. Thus, $Pv = P^*v$ for every $v \in V$ so $P = P^*$ or P is self-adjoint.

exer:7A:12

12. Let v_1, v_2 be eigenvectors of T corresponding to eigenvalues $3, 4$. Then for $e_1 = v_1/\|v_1\|, e_2 = v_2/\|v_2\|$ then $Te_1 = 4e_1$ and $Te_2 = 4e_2$. According to theorem 9.1.7 (7A) then $\langle e_1, e_2\rangle = 0$ so $\langle Te_1, Te_2\rangle = 0$ so according to Pythagorean's theorem we find $5^2 = 3^2 + 4^2 = \|Te_1\|^2 + \|Te_2\|^2 = \|T(e_1+e_2)\|^2$ so $\|T(e_1+e_2)\| = 5$ and note that $\|e_1+e_2\|^2 = \|e_1\|^2 + \|e_2\|^2 = 2$ so $\|e_1+e_2\| = \sqrt{2}$.

exer:7A:13

13. $\mathcal{M}(T) = \begin{pmatrix} 2 & -3 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

exer:7A:14 14. Since T is normal so according to theorem 9.1.7 (7A) then $\langle Tv, Tw\rangle = \langle v, w\rangle = 0$ so $2^2(3^2+4^2) = \|Tv\|^2 + \|Tw\|^2 = \|T(v+w)\|^2$ so $\|T(v+w)\| = 10$.

exer:7A:15 15. First, we will find adjoint of T. For any $w \in V$ then

$$\langle v, T^*w\rangle = \langle Tv, w\rangle = \langle v, u\rangle \langle x, w\rangle = \langle v, \langle w, x\rangle u\rangle .$$

Thus, $T^*w = \langle w, x\rangle u$.

(a) With $\mathbf{F} = \mathbf{R}$, T is self-adjoint then $Tv = T^*v$ for every $v \in V$ so $\langle v, u\rangle x = \langle v, x\rangle u$ so u, x is linearly dependent.

If u, x is linearly dependent with $u = \lambda x$ then for any $v \in V$ we have $\langle v, u\rangle x = \langle v, \lambda x\rangle u/\lambda = \langle v, x\rangle u$ so $T = T^*$ or self-adjoint.

(b) We have $TT^*v = \langle v, x\rangle Tu = \langle v, x\rangle \|u\|^2 x$ and $T^*Tv = \langle v, u\rangle T^*x = \langle v, u\rangle \|x\|^2 u$. If T is normal then $TT^* = T^*T$ then u, x is linearly dependent.

If $u = \lambda x$ then for every $v \in V$, we have

$$\begin{aligned}
\langle v, x\rangle \|u\|^2 x &= \left\langle v, \lambda^{-1} u\right\rangle |\lambda|^2 \|x\|^2 x, \\
&= \left\langle v, \overline{\lambda}/|\lambda|^2 u\right\rangle |\lambda|^2 \|x\|^2 x, \\
&= \lambda/|\lambda|^2 \langle v, u\rangle |\lambda|^2 \|x\|^2 x, \\
&= \langle v, u\rangle \|x\|^2 u.
\end{aligned}$$

This follows $TT^*v = T^*Tv$ for every $v \in V$ so T is normal.

exer:7A:16 16. First, we will show that $\operatorname{null} T = \operatorname{null} T^*$. Indeed, we have $v \in \operatorname{null} T$ iff $0 = \|Tv\|^2 = \|T^*v\|^2$ iff $T^*v = 0$ or $v \in \operatorname{null} T^*$. Thus, $\operatorname{null} T = \operatorname{null} T^*$ so $(\operatorname{null} T)^\perp = (\operatorname{null} T^*)^\perp$ so $\operatorname{range} T^* = \operatorname{range} T$ according to theorem 9.1.3 (7A).

exer:7A:17 17. We show that for any positive integer k then $\operatorname{null} T^k = \operatorname{null} T^{k+1}$. Indeed, it's obvious that if $v \in \operatorname{null} T^k$ then $v \in \operatorname{null} T^{k+1}$. If $v \in \operatorname{null} T^{k+1}$ then $T^k v \in \operatorname{null} T$ and since $\operatorname{null} T = \operatorname{null} T^*$ according to exercise 16 (7A) so $T^k v \in \operatorname{null} T^*$. On the other hand, we also have $T^k v \in \operatorname{range} T$ and $\operatorname{null} T^* = (\operatorname{range} T)^\perp$ so we must have $\|T^k v\|^2 = 0$ so $T^k v = 0$, i.e. $v \in \operatorname{null} T^k$. Thus, we obtain $\operatorname{null} T^k = \operatorname{null} T^{k+1}$ for every positive integer k.

Note that if T is normal then T^k is also normal for every positive integer k. Hence, $\operatorname{null} T^k = \operatorname{null} (T^k)^* = (\operatorname{range} T^k)^\perp$ and similarly, $\operatorname{null} T^{k+1} = (\operatorname{range} T^{k+1})^\perp$. Thus, $\operatorname{range} T^k = \operatorname{range} T^{k+1}$ for every positive integer $k \geq 1$.

exer:7A:18 18. Consider 2-dimensional vector space V over $\mathbf{C}$ so with respect to orthonormal basis e_1, e_2 of V then $\mathcal{M}(T) = \begin{pmatrix} 1 & 1-i \\ 1+i & 1 \end{pmatrix}$ and $\mathcal{M}(T^*) = \begin{pmatrix} 1 & 1+i \\ 1-i & 1 \end{pmatrix}$. Then we have $\|Te_1\| = \sqrt{1^2 + |1+i|^2} = \sqrt{1^2 + |1-i|^2} = \|T^*e_1\|$ and similarly $\|Te_2\| = \|T^*e_2\|$. However, T is not normal.

exer:7A:19 19. Since $T(1,1,1) = 2(1,1,1)$ so $(1,1,1)$ is eigenvector of T corresponding to eigenvalue 2. If $(z_1, z_2, z_3) \in \operatorname{null} T$ then (z_1, z_2, z_3) is eigenvector of T corresponding to eigenvalue 0. Hence, according to theorem 9.1.7 (7A) then $\langle (z_1, z_2, z_3), (1,1,1) \rangle = z_1 + z_2 + z_3 = 0$.

exer:7A:20 20. Note that both $\Phi_V \circ T^*$ and $T' \circ \Phi_w$ are linear map from W to V'. Hence, it suffices to prove that for every $w \in W$ then $(\Phi_V \circ T^*)(w) = (T' \circ \Phi_W)(w)$ or $\Phi_V(T^*w) = T'(\Phi_W w)$. Since there two are linear functionals on V so it again suffices to prove that for every $v \in V$ then $(\Phi_V(T^*w))v = (T'(\Phi_W w))v$. Indeed, according to definition of Φ_V, we have $(\Phi_V(T^*w))v = \langle v, T^*w \rangle$. On the other hand, $(T'(\Phi_W w))v = ((\Phi_W w) \circ T)v = (\Phi_W w)(Tv) = \langle Tv, w \rangle$. Since $\langle v, T^*w \rangle = \langle Tv, w \rangle$ so we obtain the desired equality.

exer:7A:21 21. From exercise 4 (6B), we find the list

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \ldots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \cdots, \frac{\sin nx}{\sqrt{\pi}}$$

is the orthonormal basis of V.

(a) Note that we are talking about vector space over $\mathbf{R}$ so $\mathcal{M}(D^*)$ with respect to the mentioned basis is eventually transpose of $\mathcal{M}(D)$. Hence, to show $D^* = -D$, it suffices to prove that $\mathcal{M}(D)_{j,i} + \mathcal{M}(D)_{i,j} = 0$. Note that the entry in row i, column j of $\mathcal{M}(D)$ is $\langle De_j, e_i \rangle$ where e_k is the k-th vector in the orthonormal basis mentioned above. It suffices to show $\langle De_j, e_i \rangle + \langle De_i, e_j \rangle = 0$ for every $1 \le i, j \le 2n+1$. From integrals in exercise 4 (6B), we know that if $|j - i| \neq n$ then $\langle De_i, e_j \rangle = 0$. Hence, we only need to consider when $|i - j| = n$, i.e. if $e_i = \cos(mx)/\sqrt{\pi}$ and $e_j = \sin(mx)/\sqrt{\pi}$ for some $1 \le m \le n$. We have $\langle De_i, e_j \rangle = -m \|e_j\|^2 = -m$ and $\langle De_j, e_i \rangle = m\|e_i\|^2 = m$ so $\langle De_i, e_j \rangle + \langle De_j, e_i \rangle = 0$. Thus, we obtain $D^* = -D$ as desired. Note that $\mathcal{M}(D) \neq 0$ so $D^* \neq D$ so D is not self-adjoint but D is normal.

(b) It suffices to prove $T = T^*$, i.e. $\langle Te_j, e_i \rangle = \langle Te_i, e_j \rangle$ for every $1 \le i, j \le 2n+1$. From integrals in exercise 4 (6B), when $i \neq j$ then $\langle Te_j, e_i \rangle = \langle Te_i, e_j \rangle = 0$. Thus, we obtain $T = T^*$ or T is self-adjoint.

9.3. 7B: The Spectral Theorem

theo:7.24:7B **Theorem 9.3.1** (7.24, Complex Spectral Theorem) Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Then the following are equivalent :

(a) T is normal.

(b) V has an orthonormal basis consisting of eigenvectors of T.

(c) T has a diagonal matrix with respect to some orthonormal basis of V.

theo:7.29:7B **Theorem 9.3.2** (7.29, Real Spectral Theorem) Suppose $\mathbf{F} = \mathbf{R}$ and $T \in \mathcal{L}(V)$. Then the following are equivalent:

(a) T is self-adjoint.

(b) V has an orthonormal basis consisting of eigenvectors of T.

(c) T has a diagonal matrix with respect to some orthonormal basis of V.

How to find orthonormal basis consisting of eigenvectors of a normal/self-adjoint operator T:

- Solve $Tv = \lambda v$ to find all eigenvalues.
- Find a basis of $E(\lambda, T)$ for each eigenvalue λ of T.
- Using Gram-Schidmt Procedure 8.3.2 (6B), find orthonormal basis of $E(\lambda, T)$ for each eigenvalue λ of T.
- Combines all the bases of $E(\lambda, T)$, given that T is self-adjoint/normal, we will obtain a orthonormal basis consisting of eigenvectors of T. This is true according to two exercises 4, 5 (7B).

9.4. Exercises 7B

exer:7B:1 1. Pick $T \in \mathcal{L}(\mathbf{R}^3)$ so $T(1,1,0) = 2(1,1,0), T(0,1,0) = 3(0,1,0)$ and $T(0,0,1) = 4(0,0,1)$ then with respect to standard basis (which is also orthonormal basis with respect to usual inner product) we have $\mathcal{M}(T) = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ so $\mathcal{M}(T^*) = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$. Thus T is not self-adjoint.

exer:7B:2 2. Since T is self-adjoint so according to Real/Complex Spectral Theorem, T has an diagonal matrix so $V = E(2,T) \oplus E(3,T)$. So every v can be represented as $u + w$ where $u \in E(2,T), w \in E(3,T)$. Hence $(T^2 - 5T + 6I)v = (T - 3I)(T - 2I)(u + w) = (T - 3I)[(T - 2I)u + (T - 2I)w] = (T - 3I)(T - 2I)w = (T - 2I)[(T - 3I)w] = 0$. Thus, $T^2 - 5T + 6I = 0$.

exer:7B:3 3. Pick $T \in \mathcal{L}(\mathbf{C})^3)$ so $T(1,0,0) = 3(1,0,0), T(0,1,0) = 2(0,1,0)$ and $T(0,0,1) = 2(1,1,1)$. Then $(T^2 - 5T + 6I)(0,0,1) = (-2,0,0)$ so $T^2 - 5T + 6I \neq 0$. 2 and 3 are the only eigenvalues of T.

exer:7B:4 4. If T is normal then according to theorem 9.1.7 (7A), every pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal. According to Complex Spectral Theorem 9.3.1 then T has diagonal matrix so from theorem 7.5.2 (5C), $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$.

Conversely, according to theorem 8.3.3 (6B), each finite-dimensional vector space $E(\lambda_i, T)$ has an orthonormal basis. Combining all orthonormal bases of vector spaces $E(\lambda_i, T)$ for $1 \leq i \leq m$ we obtain a new orthonormal basis (since every vector in $E(\lambda_i, T)$ is orthogonal to every vector in $E(\lambda_j, T)$ for $j \neq i$). This orthonormal basis has length dim $E(\lambda_1, T) +$

$\ldots + \dim E(\lambda_m, T)$, which equals to $\dim V$ because $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$. Thus, we obtain an orthonormal basis of V consisting of eigenvectors of T so T is normal.

exer:7B:5 5. Completely the same as previous exercise 4, except this time we use Real Spectral Theorem instead of Complex Spectral Theorem and note that any self-adjoint operator is normal.

exer:7B:6 6. One direction is already proven by theorem 9.1.5 (7A). Thus, if every eigenvalue of a normal operator T is real then according to Complex Spectral Theorem 9.3.1 (7B), T has diagonal matrix $\mathcal{M}(T)$ consisting of all real numbers. With this and theorem 9.1.4 (7A), we find $\mathcal{M}(T) = \mathcal{M}(T^*)$ so T is self-adjoint.

exer:7B:7 7. Since T is normal so according to exercise 16, 17 (7A) then $=$ range $T^* =$ range $T =$ range T^7 which follows $V = \text{null } T^7 \oplus \text{range } T^7$ from theorem 9.1.3 (7A) and 8.5.1 (6C). Thus, each $v \in V$ can be written as $v = T^7 u + w$ where $w \in \text{null } T^7$. Thus, $T^2 v = T^2(T^7 u + w) = T^9 u = T^8 u = T(T^7 u + w) = Tv$. We conclude $T^2 = T$. If $v \neq 0$ is an eigenvector of T corresponding to eigenvalue $\lambda \in \mathbf{C}$ then $T^v = \lambda^2 v = \lambda v = Tv$ so $\lambda(\lambda - 1)v = 0$ so $\lambda \in \{0, 1\}$. Thus, all eigenvalues of normal operator T are real so according to exercise 6 (7B), T is self-adjoint.

exer:7B:8 8. UNSOLVEd

exer:7B:9 9. Every normal operator T has a diagonal matrix according to Complex Spectral Theorem 9.3.1 (7B). Pick operator $S \in \mathcal{L}(V)$ so S has diagonal matrix $\mathcal{M}(S)$ and $\mathcal{M}(S)_{i,i} = \sqrt{\mathcal{M}(T)_{i,i}}$ for every $1 \le i \le n$ (this is possible because we are considering complex vector space). Hence, $\mathcal{M}(S^2) = (\mathcal{M}(S))^2 = \mathcal{M}(T)$ so $S^2 = T$.

exer:7B:10 10. Theorem 7.3.5 (look at the proof) makes the construction easier. It suffices to choose an operator T so that T has no eigenvalue. Pick $V = \mathbf{R}^2$ and $T(x, y) = (-y, x)$ then if $T(x, y) = \lambda(x, y)$ we find $-y = \lambda x, x = \lambda y$ so $y(\lambda^2 + 1) = 0$, which happens when $y = 0$ so $x = 0$. Thus, T has no eigenvalues. Hence, since for any $v \in \mathbf{R}^2, v \neq 0$ then $T^2 v, Tv, v$ is linearly dependent so there exists real numbers b, c so $b^2 < 4c$ and $T^2 v + bTv + cv = 0$. The condition $b^2 < 4c$ guarantees that T has no eigenvalue, otherwise we can factorise $T^2 v + bTv + cv$ and then obtain eigenvalue for T.

exer:7B:11 11. If T is self-adjoint then T is also normal so according to Real/Complex Spectral Theorem, T has a diagonal matrix $\mathcal{M}(T)$ with respect to some orthonormal basis of V. Pick $S \in \mathcal{L}(V)$ so S also has a diagonal matrix so $\mathcal{M}(S)_{i,i} = \sqrt[3]{\mathcal{M}(T)_{i,i}}$ for every i. We can conclude $S^3 = T$ from here.

exer:7B:12 12. Since T is self-adjoint so according to Complex/Real Spetral Theorem, V has an orthonormal basis consisting of eigenvectors of T, i.e. each $v \in V, \|v\| = 1$ can be represented as $v = \sum_{i=1}^m a_i u_i$ where $u_i \in E(\lambda_i, T)$ and $\langle u_i, u_j \rangle = 0$ for $1 \le i < j \le m$ and $a_i \in \mathbf{F}$ so $\sum_{i=1}^m |a_i|^2 = 1$. Hence,

$$\|Tv - \lambda v\|^2 = \left\| \sum_{i=1}^m (\lambda_i - \lambda) a_i u_i \right\|^2 = \sum_{i=1}^m |(\lambda_i - \lambda) a_i|^2.$$

If $|\lambda_i - \lambda| \geq \epsilon$ for all $1 \leq i \leq m$ then $\|Tv - \lambda v\|^2 \geq \epsilon^2 \sum_{i=1}^m |a_i|^2 = \epsilon^2$ so $\|Tv - \lambda v\| \geq \epsilon$, a contradiction. Thus, there must exists i so $|\lambda_i - \lambda| < \epsilon$.

exer:7B:13 13. According to exercise 3 (7A), if U is invariant under T then $U^\perp$ is invariant under T^*. Hence, if T is normal then $T^*|_{U^\perp} \in \mathcal{L}(U^\perp)$ is normal.

We prove Complex Spectral Theorem by induction on dim V. In a complex inner product space V, there always exists an eigenvector $u \neq 0$ corresponding to eigenvalue λ of T. Let $U = \text{span } (u)$ then since dim $U^\perp <$ dim V, there exists an orthonormal basis $e_1, \ldots, e_{n-1}$ of $U^\perp$ consisting of eigenvectors of $T^*|_{U^\perp}$. Since T is normal so according to exercise 2 (7A), $e_1, \ldots, e_{n-1}$ is also orthonormal basis consisting of eigenvectors of T. Note that eigenvalues of T of these eigenvectors must be different from λ, otherwise $u \in U^\perp$ which means $u = 0$, a contradiction. Thus, V has orthonormal basis $u, e_1, \ldots, e_{n-1}$ consisting of eigenvectors of T.

exer:7B:14 14. One direction is true according to Real Spectral Theorem. If U has basis $u_1, \ldots, u_n$ consisting of eigenvectors of T, consider inner product on U so $\langle u_i, u_j \rangle = 0$ for all $1 \leq i < j \leq n$ then this inner product makes T into a self-adjoint operator according to Real Spectral Theorem.

exer:7B:15 15. It suffices to find $a \in \mathbf{R}$ so

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

9.5. 7C: Positive Operators and Isometries

Definition 9.5.1. An operator $T \in \mathcal{L}(V)$ is called **positive** if T is self-adjoint and $\langle Tv, v \rangle \geq 0$ for all $v \in V$.

Another terminology for the above definition is **positive semidefinite**. A positive semidefinite operator that satisfies $\langle Tv, v \rangle > 0$ for all non-zero v is called **positive definite**.

theo:7.35:7C **Theorem 9.5.2** (7.35, Positive Operators) Let $T \in \mathcal{L}(V)$. Then the following are equivalent:

(a) T is positive;

(b) T is self-adjoint and all the eigenvalues of T are nonnegative;

(c) T has positive square root;

(d) T has self-adjoint square root;

(e) There exists an operator $R \in \mathcal{L}(V)$ such that $T = R^*R$.

theo:7.42:7C **Theorem 9.5.3** (7.42, Isometries) Suppose $S \in \mathcal{L}(V)$. Then the following are equivalent:

(a) S is an isometry;

(b) $\langle Su, Sv\rangle = \langle u, v\rangle$ for all $u, v \in V$;

(c) $Se_1, \dots, Se_n$ is orthonormal for every orthonormal list of vectors $e_1, \dots, e_n$ in V;

(d) there exists an orthonormal basis $e_1, \dots, e_n$ of V such that $Se_1, \dots, Se_n$ is orthonormal;

(e) $S^*S = I$;

(f) $SS^* = I$;

(g) S^* is an isometry;

(h) S is invertible and $S^{-1} = S^*$.

theo:7.43:7C **Theorem 9.5.4** (7.43) Suppose V is a complex inner product space and $S \in \mathcal{L}(V)$. Then the following are equivalent:

(a) S is isometry.

(b) There is an orthonormal basis of V consisting of eigenvectors of S whose corresponding eigenvalues all have absolute value 1.

9.6. Exercises 7C

exer:7C:1 1. Let e_1, e_2 be orthonormal basis of V. Let $Te_1 = xe_1 + ye_2$ and $Te_2 = ye_1 + ze_2$ with $x, y, z \in \mathbf{R}$ so $2y + x + z < 0, x, z \geq 0$. Hence, T is self-adjoint and $\langle Te_1, e_1\rangle = x \geq 0, \langle Te_2, e_2\rangle = z \geq 0$. On the other hand,

$$\langle T(e_1 + e_2), e_1 + e_2\rangle = \langle (x + y)e_1 + (y + z)e_2, e_1 + e_2\rangle = x + 2y + z < 0.$$

Thus, T is not positive.

exer:7C:2 2. We have

$$\langle T(v - w), v - w\rangle = \langle w - v, v - w\rangle = -\|v - w\|^2 \geq 0.$$

Thus, we must have $\|v - w\|^2 = 0$ or $v = w$.

exer:7C:3 3. We have for every $u \in U$ then $\langle T|_U u, u\rangle = \langle Tu, u\rangle \geq 0$ so $T|_U$ is positive operator on U.

exer:7C:4 4. For every $v \in V$, we have $\langle T^*Tv, v\rangle = \langle Tv, Tv\rangle \geq 0$ and $(T^*T)^* = (T^*T$ so T^*T is positive operator on V. Similarly, TT^* is positive operator on V.

exer:7C:5 5. If T, S is positive operators on V then for any $v \in V$ we have $\langle (T + S)v, v\rangle = \langle Tv, v\rangle + \langle Sv, v\rangle \geq 0$ and $(T + S)^* = T^* + S^* = T + S$ so $T + S$ is positive operator on V.

exer:7C:6 6. If T is positive operator on V then according to theorem 9.5.2 (7C), T has self-adjoint square root $R \in \mathcal{L}(V)$. Hence, we have $R^k = (R^*)^k = (R^k)^*$ so R^k is self-adjoint for every positive integer k. Hence, $T^k = (R^2)^k = (R^k)^2$ so again from theorem 9.5.2 (7C), T^k is a positive operator on V.

exer:7C:7 7. If $\langle Tv, v\rangle > 0$ for every $v \in V, v \neq 0$ then if T is invertible (otherwise there exists $v \in V, v \neq 0$ so $Tv = 0$ which leads to $\langle Tv, v\rangle = 0$, a contradiction).

   Conversely, if T is invertible. Since T is positive so according to theorem 9.5.2 (7C), there exists operator $R \in \mathcal{L}(V)$ so $R^*R = T$. Hence, $\langle Tv, v\rangle = \|Rv\|^2$. Since T is invertible so $Rv \neq 0$ for all $v \in V$, otherwise $Tv = R^*Rv = 0$ a contradiction. Thus, $\|Rv\|^2 > 0$ or $\langle Tv, v\rangle > 0$ for every $v \in V, v \neq 0$.

exer:7C:8 8. It suffices to check the definiteness condition of inner product. We have $\langle v, v\rangle_T = 0$ iff $\langle Tv, v\rangle = 0$. Applying previous exercise 7 (7C), the condition $\langle Tv, v\rangle = 0$ for only $v = 0$ is equivalent to T being invertible.

exer:7C:9 9. Let S be a self-adjoint square root of identity operator I then according to Real/Complex Spectral Theorem 9.3.1 (7B), V has an orthonormal basis e_1, e_2 consisting of eigenvectors of S corresponding to eigenvalues λ_1, λ_2. Thus, we have $S^2 e_1 = \lambda_1^2 e_1 = Ie_1 = e_1$ so $\lambda_1^2 = 1$. Similarly, $\lambda_2^2 = 1$. Hence, identity operator on $\mathbf{F}^2$ has finitely many self-adjoint square roots.

exer:7C:10 10. If (a) holds, i.e. S is isometry then from theorem 9.5.3 (7C), S^* is also isometry which leads to (b). If any of (b),(c) or (d) holds, we can follow that S^* is isometry which can leads back to other conditions according to theorem 9.5.3 (7C).

exer:7C:11 11. Similar to exercise 12 (5C). Since T_1 is normal so according to Complex Spectral Theorem 9.3.1 (7B), there exists an orthonormal basis v_1, v_2, v_3 consisting of eigenvectors $2, 5, 7$ of T_1. Similarly, there exists an orthonormal basis w_1, w_2, w_3 consisting of eigenvectors $2, 5, 7$ of T_2. Let $S \in \mathcal{L}(\mathbf{F}^3)$ so $Sv_1 = w_1, Sv_2 = w_2$ and $Sw_3 = w_3$ then S is invertible. Similarly to exercise 15 (5C), we can show that $T_1 = S^{-1}TS$. Now it suffices to show that S is isometry. Indeed, for any $v \in V, v = \sum_{i=1}^3 a_i e_i$ then $\|Sv\|^2 = \left\|\sum_{i=1}^3 a_i w_i\right\|^2 = \sum_{i=1}^3 |a_i|^2 = \|v\|^2$. Thus, S is isometry so according to theorem 9.5.3 (7C), $S^{-1} = S^*$ so $T_1 = S^* T_2 S$.

exer:7C:12 12. It is essentially exercise 13 (5C).

exer:7C:13 13. It is false. Pick $S \in \mathcal{L}(V)$ for $Se_j = e_j$ for all $j \neq 2$ and $Se_2 = x_1 e_1 + x_2 e_2$ where $x_1, x_2 \in \mathbf{R}$ so $|x_1|^2 + |x_2|^2 = 1$ to guarantee $\|Se_2\| = 1$. Hence $\|S(e_1 + e_2)\|^2 = \|(x_1+1)e_1 + x_2 e_1\|^2 = |x_1 + 1|^2 + |x_2|^2 = 2 + 2x_1 \neq 2 = \|e_1 + e_2\|^2$. Thus, S is not an isometry.

exer:7C:14 14. We've already shown that T is self-adjoint so $-T$ is also self-adjoint. We have

$$\langle (-T)\sin ix, \sin ix\rangle = \langle \sin ix, \sin ix\rangle > 0,\ \langle (-T)\cos ix, \cos ix\rangle = \langle \cos ix, \cos ix\rangle > 0.$$

   And $\langle (-T)e_i, e_j\rangle = 0$ for $i \neq j$ (see exercise 21 (7A) for more details). With this, we can conclude that $\langle (-T)v, v\rangle > 0$ for all $v \in V$, i.e. $-T$ is positive.

9.7. 7D: Polar Decomposition and Singular Value Decomposition

theo:7.45:7D **Theorem 9.7.1** (7.45, Polar Decomposition) Suppose $T \in \mathcal{L}(V)$. Then there exists an isometry $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$.

theo:7.51:7D **Theorem 9.7.2** (7.51 Singular Value Decomposition) Suppose $T \in \mathcal{L}(V)$ has singular values $s_1, \dots, s_n$. Then there exist orthonormal bases $e_1, \dots, e_n$ and $f_1, \dots, f_n$ of V such that

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \dots + s_n \langle v, e_n \rangle f_n$$

for every $v \in V$.

theo:7.52:7D **Theorem 9.7.3** (7.52) Suppose $T \in \mathcal{L}(V)$. Then the singular values of T are the nonnegative square roots of the eigenvalues of T^*T, with each eigenvalue λ repeated $\dim E(\lambda, T^*T)$ times.

How to find positive operator R if R^2 is known:

- Since R^2 is also positive so first is to find orthonormal basis $e_1, \dots, e_n$ of V consisting of eigenvectors of R^2 (see 7B).
- If $R^2 e_i = \lambda_i e_i$ then $Re_i = \sqrt{\lambda_i} e_i$. Hence,

$$Rv = R\left(\sum_{i=1}^{n} \langle v, e_i \rangle e_i\right) = \sum_{i=1}^{n} \langle v, e_i \rangle \sqrt{\lambda_i} e_i.$$

How to find singular values of an operator T:

- Find T^* (see 7A) then find T^*T then find all eigenvalues of T^*T. Hence, singular values of T are nonnegative square roots of the eigenvalues of T^*T according to theorem 9.7.3 (7D).
- If T is self-adjoint: The singular values of T equals the absolute values of the eigenvalues of T according to exercise 10 (7D).

9.8. Exercises 7D

exer:7D:1 1. According to solution in exercise 15 (7A) then $T^*w = \langle w, x \rangle u$ so $T^*Tv = \langle v, u \rangle T^*x = \langle v, u \rangle \langle x, x \rangle u$. One can verify that $\sqrt{T^*T}v = \frac{\|x\|}{\|u\|} \langle v, u \rangle u$ is indeed true.

exer:7D:2 2. Let $T(a, b) = (5b, 0)$ then $T^*(x, y) = (0, 5x)$. Hence, $T^*T(x, y) = (0, 25y)$. Hence, $\sqrt{T^*T}(x, y) = (0, 5y)$ which has two singular values $0, 5$ while 0 is the only eigenvalue of T.

exer:7D:3 3. Apply Polar Decomposition 9.7.1 (7D) to T^* we find there exists an isometry $S \in \mathcal{L}(V)$ so $T^* = S\sqrt{(T^*)^*T^*} = S\sqrt{TT^*}$. Hence, $T = (T^*)^* = \left(\sqrt{TT^*}\right)^* S^* = \sqrt{TT^*}S^*$ because $\sqrt{TT^*}$ is positive. Note that S is an isometry so S^* is also an isometry.

exer:7D:4 4. Since s is singular value of T so there exists $v \in V, \|v\| = 1$ so $\sqrt{T^*T}v = sv$. Hence, we have $\|Tv\| = \|\sqrt{T^*T}v\| = |s|\|v\| = s$.

exer:7D:5 5. We have

$$\langle (a,b), T^*(x,y)\rangle = \langle T(a,b),(x,y)\rangle = \langle (-4b,a),(x,y)\rangle = -4b\overline{x} + a\overline{y} = \langle (a,b),(y,-4x)\rangle .$$

Hence, $T^*(x,y) = (y,-4x)$. Hence, $T^*T(x,y) = T^*(-4y,x) = (x,16y)$. Thus, singular values of T are $1,4$.

exer:7D:6 6. We have for all $a_2, a_1, a_0 \in \mathbf{R}$ then

$$\begin{aligned}
\left\langle a_2x^2 + a_1x + a_0, T^*(b_2x^2 + b_1x + b_0)\right\rangle &= \left\langle T(a_2x^2 + a_1x + a_0), b_2x^2 + b_1x + b_0\right\rangle ,\\
&= \left\langle 2a_2x + a_1, b_2x^2 + b_1x + b_0\right\rangle ,\\
&= \int_{-1}^{1} (2a_2x + a_1)(b_2x^2 + b_1x + b_0)\, dx,\\
&= \frac{2}{3}(2a_2b_1 + a_1b_2) + 2a_1b_0,\\
&= a_2 \cdot \frac{4}{3}b_1 + a_1\left(\frac{2}{3}b_2 + 2b_0\right).
\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
\left\langle a_1x^2 + a_1x + a_0, c_2x^2 + c_1x + c_0\right\rangle &= 2a_0c_0 + \frac{2}{3}(a_1c_1 + a_2c_0 + a_0c_2) + \frac{2}{5}a_2c_2,\\
&= a_2\left(\frac{2}{5}c_2 + \frac{2}{3}c_0\right) + a_1 \cdot \frac{2}{3}c_1 + 2a_0c_0.
\end{aligned}$$

Thus, we must have $c_0 = 0, \frac{4}{3}b_1 = \frac{2}{5}c_2 + \frac{2}{3}c_0$ and $\frac{2}{3}b_1 + 2b_0 = \frac{2}{3}c_1$. This follows $c_0 = 0, c_2 = \frac{10}{3}b_1$ and $c_1 = b_1 + 3b_0$. We obtain $T^*(b_2x^2 + b_1x + b_0) = \frac{10}{3}b_1x^2 + (b_1 + 3b_0)x$. Thus,

$$T^*T(a_2x^2 + a_1x + a_0) = T^*(2a_2x + a_1) = (2a_2 + 3a_1)x.$$

Thus, T^*T has eigenvalues 3 which repeated 1 time and 0 which repeated 2 times. Thus, singular values of T are $0, 0, \sqrt{3}$ according to theorem 9.7.3 (7D).

exer:7D:7 7. We have

$$\begin{aligned}
\langle (z_1,z_2,z_3), T^*(x_1,x_2,x_3)\rangle &= \langle T(z_1,z_2,z_3),(x_1,x_2,x_3)\rangle ,\\
&= \langle (z_3, 2z_1, 3z_2),(x_1,x_2,x_3)\rangle ,\\
&= \langle (z_1,z_2,z_3),(2x_2,3x_3,x_1)\rangle .
\end{aligned}$$

Thus, $T^*(x_1,x_2,x_3) = (2x_2,3x_3,x_1)$. Hence, $T^*T(z_1,z_2,z_3) = T^*(z_3,2z_1,3z_2) = (4z_1,9z_2,z_3)$. Thus, $\sqrt{T^*T}(z_1,z_2,z_3) = (2z_1,3z_2,z_3)$. Let $S \in \mathcal{L}(V)$ so $S(2z_1,3z_2,z_3) = (z_3,2z_1,3z_2)$ then S is an isometry.

exer:7D:8 8. Since $T = SR$ so $T^* = (SR)^* = R^*S^* = RS^*$ so $T^*T = RS^*SR = R^2$ so $R = \sqrt{T^*T}$.

exer:7D:9 9. Since $\|Tv\| = \|\sqrt{T^*T}v\|$ so T is invertible iff $\sqrt{T^*T}$ is invertible iff null $\sqrt{T^*T} = \{0\}$ iff range $\sqrt{T^*T} = V$. Thus, it suffices to show that range $T^*T = V$ iff there exists a unique isometry $S \in \mathcal{L}(V)$ so $T = S\sqrt{T^*T}$.

   Indeed, if range $\sqrt{T^*T} = V$ then obviously isometry operator S is uniquely determined since S must satisfy $T = S\sqrt{T^*T}$. Conversely, if there exists a unique isometry $S \in \mathcal{L}(V)$ so $T = S\sqrt{T^*T}$ but range $\sqrt{T^*T} \neq V$, which means range $\sqrt{T^*T})^\perp \neq \{0\}$. Hence, by looking back proof of Polar Decomposition 9.7.1 (7D) from the book, for $m \geq 1$, let $e_1, \ldots, e_m$ and $f_1, \ldots, f_m$ be orthonormal bases of (range $\sqrt{T^*T})^\perp$ and (range $T)^\perp$, respectively. Hence, we have at least two ways to choose linear map $S_2 :$ (range $\sqrt{T^*T})^\perp \to$ (range $T)^\perp$, one is $S_2(a_1e_1 + \ldots a_me_m) = a_1f_1 + \ldots + a_mf_m$ or $S_2(a_1e_1 + \ldots, a_mf_m) = -a_1f_1 - \ldots - a_mf_m$. This follows there are at least two possible isometries S so $T = S\sqrt{T^*T}$, a contradiction. Thus, we must have range $\sqrt{T^*T} = V$ or T is invertible.

exer:7D:10 10. According to theorem 9.7.3 (7D), the singular values of T are the nonnegative square roots of the eigenvalues of T^*T. Since T is self-adjoint so $T^*T = T^2$ so eigenvalues of T are the nonnegative square roots of the eigenvalues of T^2. Thus, the singular values of T equal the absolute values of the eigenvalues of T, repeated appropriately.

exer:7D:11 11. It suffices to show that T^*T and TT^* have the same eigenvalues values that are repeated under same number of times. Since T^*T is self-adjoint so according to Spectral Theorem, there exists an orthonormal basis $e_1, \ldots, e_n$ consisting of eigenvectors of T^*T. WLOG, among these vectors, say $f_1, \ldots, f_m$ are eigenvectors of T^*T corresponding to eigenvalue λ. Note that $m = \dim E(\lambda, T^*T)$.

    One the other hand, $TT^*(Tf_i) = T(T^*Tf_i) = \lambda Tf_i$ so $Tf_1, \ldots, Tf_m$ are eigenvectors of TT^* corresponding to eigenvalues λ. For any $1 \leq i < j \leq m$ then $\langle Tf_i, Tf_j\rangle = \langle f_i, T^*Tf_j\rangle = \lambda \langle f_i, f_j\rangle = 0$ so $Tf_1, \ldots, Tf_m$ is also an orthonormal list which means $Tf_1, \ldots, Tf_m$ is linearly independent. This follows that eigenvalue λ under operator TT^* is repeated at least $\dim E(\lambda, T^*T)$ times, i.e. $\dim E(\lambda, TT^*) \geq \dim E(\lambda, T^*T)$. Similarly and we obtain $\dim E(\lambda, TT^*) = \dim E(\lambda, T^*T)$ and can conclude that TT^* and T^*T have same eigenvalues that are repeated same number of times. Thus, T and T^* have same singular values.

exer:7D:12 12. The statement is false. According to theorem 9.7.3 (7D), the singular values of T^2 are the nonnegative square roots of the eigenvalues of $(T^2)^*T^2 = (T^*)^2T^2$ and the singular values of T are the nonnegative square roots of the eigenvalues of T^*T. Hence, the singular values of T^2 equal the squares of the singular values of T when the eigenvalues of $(T^*)^2T^2$ equals the squares of the eigenvalues of T^*T. Note that $(T^*)^2T^2$ and T^*T are self-adjoint operators so we must have $(T^*T)^2 = (T^*)^2T^2$, which is incorrect when T is not normal. Choose a nonnormal operator T in 2-dimensional vector space V as a counterexample.

exer:7D:13 13. Since for all $v \in V$ then $\|Tv\| = \|\sqrt{T^*T}v\|$ so T is invertible iff $\sqrt{T^*T}$ is invertible iff 0 is not an eigenvalue of $\sqrt{T^*T}$ iff 0 is not s singular value of T.

exer:7D:14 14. Again, from $\|Tv\| = \|\sqrt{T^*T}v\|$ so dim range $T =$ dim range $\sqrt{T^*T}$ and dim range $\sqrt{T^*T} = n -$ dim $E(0, \sqrt{T^*T})$ from the fact that $\sqrt{T^*T}$ is positive so $\sqrt{T^*T}$ is diagonalizable. Note that according to the definition of singular values, dim $E(0, \sqrt{T^*T})$ is number of repetition of singular value 0 of T. Thus, dim range $T = n -$ dim $E(0, \sqrt{T^*T})$ equals the number of nonzero singular values of T.

exer:7D:15 15. If S has singular values $s_1, \ldots, s_n$ then according to theorem 9.7.2 (7D), there exist orthonormal bases $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ of V such that $Sv = s_1 \langle v, e_1 \rangle f_1 + \ldots + s_n \langle v, e_n \rangle f_n$ for every $v \in V$. This follows for every $v \in V$, $\|Sv\|^2 = \sum_{i=1}^n |s_i \langle v, e_i \rangle|^2$. Hence, if all singular values of S equal 1 then $\|Sv\| = \|v\|$ so S is an isometry. Conversely, if S is an isometry then $\|Se_i\| = s_i = 1$ for all $1 \le i \le n$ so all singular values of S equal 1.

exer:7D:16 16. Note that with $T_1 = S_1T_2S_2$ then $T_1^*T = S_2^*(T_2^*T_2)S_2$. The idea to solve this problem is similar to exercise 11 (7C).

If T_1, T_2 have same singular values then two positive operators $T_1^*T_1$ and $T_2^*T_2$ have same eigenvalues $\lambda_1, \ldots, \lambda_n$. According to Spectral Theorem, let $e_1, \ldots, e_n$ be orthonormal basis consisting of eigenvectors of $T_1^*T_1$ and $f_1, \ldots, f_n$ be orthonormal basis consisting of eigenvectors of $T_2^*T_2$. Let $S_2 \in \mathcal{L}(V)$ so $S_2e_i = f_i$ for all $1 \le i \le n$ then similarly to exercise 11 (7C), we can show that S_2 is an isometry and that $T_1^*T_1 = S_2^*(T_2^*T_2)S_2$. This leads to $\|T_1v\| = \|T_2S_2v\|$ for every $v \in V$. With this, we can construct isometry S_1 similarly to proof of Polar Decomposition 9.7.1 (7D) from the book so that $S_1T_2S_2 = T_1$.

Conversely, if there exists isometries S_1, S_2 so $S_1T_2S_2 = T_1$ then $T_1^*T_1 = S_2^*T_2^*T_2S_2$. If $f_1, f_2, \ldots, f_n$ is basis of V and are also eigenvectors of $T_2^*T_2$ corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$. Since S_2 is invertible, for each f_i, there exists $e_i \in V$ so $S_2e_i = f_i$. This follows $e_1, \ldots, e_n$ is basis of V. Hence,

$$T_1^*T_1e_i = S_2^*T_2^*T_2S_2e_i = S_2^*T_2^*T_2f_i = \lambda S_2^*f_i = \lambda_ie_i.$$

This follows $T_1^*T_1$ and $T_2^*T_2$ have same eigenvalues so T_1 and T_2 have same singular values.

exer:7D:17 17. (a) We have

$$\begin{aligned}
\langle Tv, u \rangle &= \langle s_1 \langle v, e_1 \rangle f_1 + \ldots + s_n \langle v, e_n \rangle f_n, \langle u, f_1 \rangle f_1 + \ldots + \langle u, f_n \rangle f_n \rangle, \\
&= s_1 \langle v, e_1 \rangle \langle u, f_1 \rangle + \ldots + s_n \langle v, e_n \rangle \langle u, f_n \rangle, \\
&= \langle v, s_1 \langle u, f_1 \rangle e_1 + \ldots + s_n \langle u, f_n \rangle e_n \rangle.
\end{aligned}$$

This follows $T^*u = s_1 \langle u, f_1 \rangle e_1 + \ldots + s_n \langle u, f_n \rangle e_n$.

(b) We have

$$\begin{aligned}
T^*Tv &= T^* \left(s_1 \langle v, e_1 \rangle f_1 + \ldots + s_n \langle v, e_n \rangle f_n \right), \\
&= \sum_{i=1}^n s_i \langle v, e_i \rangle T^* f_i, \\
&= s_1^2 \langle v, e_1 \rangle e_1 + \ldots + s_n^2 \langle v, e_n \rangle e_n.
\end{aligned}$$

(c) We have $\sqrt{T^*T}e_i = s_ie_i$ so

$$\begin{aligned}\sqrt{T^*T}v &= \sqrt{T^*T}\left(\langle v,e_1\rangle e_1+\ldots+\langle v,e_n\rangle e_n\right),\\ &= s_1\langle v,e_1\rangle e_1+\ldots+s_n\langle v,e_n\rangle e_n.\end{aligned}$$

(d) We have $Te_i = s_if_i$ so $T^{-1}f_i = \frac{e_i}{s_i}$. Hence,

$$\begin{aligned}T^{-1}v &= T^{-1}\left(\langle v,f_1\rangle f_1+\ldots+\langle v,f_n\rangle f_n\right),\\ &= \sum_{i=1}^n \langle v,f_i\rangle T^{-1}f_i,\\ &= \frac{\langle v,f_1\rangle e_1}{s_1}+\ldots+\frac{\langle v,f_n\rangle e_n}{s_n}.\end{aligned}$$

exer:7D:18 18. (a) If $s_1\le s_2\le\cdots\le s_n$ are all singular values of T then we have $\|Tv\|^2 = \sum_{i=1}^n s_i^2|\langle v,e_i\rangle|^2$ and $\|v\|^2 = \sum_{i=1}^n|\langle v,e_i\rangle|^2$ so it can be easily seen that $s_1\|v\|\le\|Tv\|\le s_n\|v\|$.

(b) Applying (a) we obtain $s_1\le|\lambda|\le s_n$.

exer:7D:19 19. It suffices to show that for every $\varepsilon>0$ there exists $\delta>0$ so for all $\|u-v\|<\delta$ then $\|T(u-v)\|<\varepsilon$. Indeed, if s the largest singular value of T then from previous exercise 18 (7D), we find $\|T(u-v)\|\le s\|u-v\|$. Hence, if we choose $\delta=\varepsilon/s$, we are done. Thus, T is uniformly continuous with respect to the metric d on V.

exer:7D:20 20. According to exercise 18 (7D), for any $v\in V$, we have $\|(S+T)v\|\le\|Sv\|+\|Tv\|\le(t+s)\|v\|$. On the other hand, from exercise 4 (7D), there exists $v\in V,\|v\|=1$ so $\|(S+T)v\|=r$ so this follows $r\le t+s$.

10. Chapter 8: Operators on Complex Vector Spaces

In this chapter, V denotes a finite-dimensional nonzero vector space over $\mathbf{F}$.

10.1. 8A: Generalized Eigenvectors and NilpotentOperators

theo:8.2:8A **Theorem 10.1.1** (8.2) Suppose $T \in \mathcal{L}(V)$. Then

$$\{0\} = \text{null } T^0 \subset \text{null } T^1 \subset \cdots \subset \text{null } T^k \subset \text{null } T^{k+1} \subset \cdots$$

theo:8.3:8A **Theorem 10.1.2** (8.3) Suppose $T \in \mathcal{L}(V)$. Suppose m is a nonnegative integer such that null $T^m =$ null T^{m+1}. Then

$$\text{null } T^m = \text{null } T^{m+1} = \text{null } T^{m+2} = \text{null } T^{m+3} = \ldots$$

theo:8.4:8A **Theorem 10.1.3** (8.4) Suppose $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then null $T^k =$ null T^{k+1} for all positive integer $k \geq n$.

theo:8.5:8A **Theorem 10.1.4** (8.5) Suppose $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then $V = \text{null } T^n \oplus \text{range } T^n$.

theo:8.13:8A **Theorem 10.1.5** (8.13) Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of T and $v_1, \ldots, v_m$ are corresponding generalized eigenvectors. Then $v_1, \ldots, v_m$ is linearly independent.

theo:8.18:8A **Theorem 10.1.6** (8.18) Suppose $N \in \mathcal{L}(V)$ is nilpotent. Then $N^{\dim V} = 0$.

theo:8.19:8A **Theorem 10.1.7** (8.19) Suppose N is a nilpotent operator on V. Then there exists a basis of V with respect to which the matrix of N is an upper-triangular matrix with only 0's in the diagonal.

How two determine all the generalized eigenspaces corresponding to distinct eigenvalues of operator $T \in \mathcal{L}(V)$:

1. Find all eigenvalues of T.

2. For each eigenvalue λ of T, find null space of $(T - \lambda I)^{\dim V}$, which is equal to generalized eigenspace $G(\lambda, T)$ of T corresponding to λ.

10.2. Exercises 8A

exer:8A:1

1. 0 is the only eigenvalue of T. We have $T^2(w,z) = T(z,0) = 0$ for all $(w,z) \in \mathbf{C}^2$ so $G(0,T) = \text{null } T^2 = \mathbf{C}^2$.

exer:8A:2

2. If $T(w,z) = \lambda(w,z) = (-z,w)$ then $\lambda w = -z, \lambda z = w$. This follows $\lambda^2 + 1 = 0$ so eigenvalues of T are $\lambda = \pm i$. We have
$$\begin{aligned}(T+iI)^2(w,z) &= (T+iI)(-z+iw, w+iz),\\ &= (-w-iz, -z+iw) + i(-z+iw, w+iz),\\ &= (-2w-2iz, 2iw-2z).\end{aligned}$$
Thus, $(T+iI)^2(w,z) = 0$ when $w+iz = 0$. Thus, $G(-i,T) = \text{span } (-i,1)$. We have
$$\begin{aligned}(T-iI)^2(w,z) &= (T-iI)(-z-iw, w-iz),\\ &= (iz-w, -z-iw) - i(-z-iw, w-iz),\\ &= (2iz-2w, -2z-2iw).\end{aligned}$$
Thus, $(T-iI)^2(w,z) = 0$ when $iz - w = 0$ so $G(i,T) = \text{span } (i,1)$.

exer:8A:3

3. Let $\dim V = n$. If $v \in G(\lambda,T)$ then $(T-\lambda I)^n v = 0$. Hence, $(\lambda T)^{-n}(T-\lambda I)^n v = 0$ so $(\lambda^{-1}I - T^{-1})^n v = 0$ so $v \in G(\lambda^{-1}, T^{-1})$. Similarly, if $v \in G(\lambda^{-1}, T^{-1})$ then $v \in G(\lambda, T)$. Thus, $G(\lambda,T) = G(\frac{1}{\lambda}, T^{-1})$.

exer:8A:4

4. The proof for this is completely similar to proof of theorem 10.1.5 (8A). Indeed, assume the contrary there is $v \in G(\alpha,T) \cap G(\beta,T), v \neq 0$ then let k be the largest nonnegative integer so $(T-\alpha I)^k v \neq 0$. Hence, with $w = (T-\alpha I)^k v$ we have $(T-\alpha I)w = (T-\alpha I)^{k+1} v = 0$ so $Tw = \alpha w$. This follows $(T-\lambda I)w = (\alpha-\lambda)w$ for any $\lambda \in \mathbf{F}$. Hence, we obtain, $(T-\beta I)^n w = (\alpha-\beta)^n w$. On the other hand, since $v \in G(\beta,T)$ so $(T-\beta I)^n w = (T-\alpha I)^k (T-\beta I)^n v = 0$. This follows $(\alpha-\beta)^n w = 0$ which leads to $w = 0$, a contradiction. Thus, $G(\alpha,T) \cap G(\beta,T) = \{0\}$.

exer:8A:5

5. Let $a_0, \ldots, a_{m-1} \in \mathbf{F}$ so $a_0 v + a_1 Tv + \ldots + a_{m-1} T^{m-1} v = 0$. Apply the the operator T^{m-1} to both side of this we obtain $a_0 T^{m-1} v = 0$ which leads to $a_0 = 0$. Similarly, we apply T^{m-2} to obtain $a_1 = 0$, ..., T to obtain $a_{m-2} = 0$. Thus, $a_i = 0$ for all i so $v, Tv, \ldots, T^{m-1}v$ is linearly independent.

exer:8A:6

6. Since $T(z_1,z_2,z_3) = (z_2,z_3,0)$ so $T^2 = 0$. If T has square root when $S^4 = T^2 = 0$ so S is nilpotent so according to theorem 10.1.6 (8A) then $S^2 = 0$ or $T = 0$, a contradiction. Thus, T has no square root.

exer:8A:7

7. If N is nilpotent then from theorem 10.1.6 (8A), $N^{\dim V} = 0$ or $G(0,N) = \text{null } N^{\dim V} = V$ so there exists a basis of V consisting of generalized eigenvectors corresponding to 0. Combining with exercise 4 (8A), we find 0 is the only eigenvalue of N.

exer:8A:8

8. It is false. When $V = \mathbf{C}^2$, consider to nilpotent operators $T, S \in \mathcal{L}(\mathbf{C}^2)$ so $T(x,y) = (0,x)$ and $S(x,y) = (y,0)$. Then $(T+S)^2(x,y) = (T+S)(y,x) = (x,y)$. Thus, $T+S$ is not a nilpotent operator. Thus, set of nilpotent operators on $\mathbf{C}^2$ is not a subspace of $\mathcal{L}(\mathbf{C}^2)$.

exer:8A:9 9. Since ST is nilpotent so $ST^{\dim V} = 0$ which leads to $TS^{\dim V+1} = T(ST)^{\dim V}S = 0$ so TS is nilpotent.

exer:8A:10 10. It suffices to show that $\text{null } T^{n-1} \cap \text{range } T^{n-1} = \{0\}$. Assume the contrary, if there exists $v \in \text{null } T^{n-1} \cap \text{range } T^{n-1}, v \neq 0$ then $T^{n-1}v = 0$ and there exists $u \in V, u \neq 0$ so $T^{n-1}u = v$. This follows $T^{2n-2}u = 0$ so $T^n u = 0$ according to theorem 10.1.3 (8A). Note that $v = T^{n-1}u \neq 0$ and $T^n u = 0$ so according to exercise 5 (8A), $u, Tu, \ldots, T^{n-1}u$ is linearly independent so it is basis of V. On the other hand, not that $T^i u \in \text{null } T^n$ so that means $T^n = 0$, a contradiction since T is not nilputent. Thus, we must have $\text{null } T^{n-1} \cap \text{range } T^{n-1} = \{0\}$, which leads to $V = \text{null } T^{n-1} \oplus \text{range } T^{n-1}$.

exer:8A:11 11. It is false. The idea is to pick T so T^2 doesn't have enough eigenvectors to form a basis of V. Let $T \in \mathcal{L}(\mathbf{C}^2)$ so $T(x,y) = (x, x+y)$ then $T^2(x,y) = T(x, x+y) = (x, 2x+y)$. Hence, if $\lambda(x,y) = (x, 2x+y)$ then $(\lambda - 1)x = 0$ and $(\lambda - 1)y = 2x$. With this, we obtain 1 is the only eigenvalue of T^2 and that $E(1, T^2) = \text{span } \{(0,1)\}$. This follows T^2 is not diagonalizable.

exer:8A:12 12. Let $v_1, \ldots, v_n$ be such basis of V. We prove inductively on $i \leq n$ that there exists positive integer k_i so $T^{k_i}v_i = 0$. Indeed, since $Tv_1 = 0$ so it's true for $i = 1$. If it's true for all $i \leq m < n$. Consider v_{m+1} then according to the assumption, we have $Tv_{m+1} \in \text{span } (v_1, \ldots, v_m)$ so with $k = \max\{k_1, \ldots, k_m\}$ we obtain $T^{k+1}v_{m+1} = 0$. Thus, this follows N is nilpotent.

exer:8A:13 13. Since N is nilpotent so there exists a basis of V with respect to which the matrix of N is an upper triangular matrix with only 0's in the diagonal. Therefore, from theorem 8.3.4 (look at the proof of it, because only apply the theorem itself doesn't help), there exists an orthonormal basis of V with respect to which the matrix of N is an upper triangular matrix with only 0's in the diagonal. With this and the from the proof of Complex Spectral Theorem 9.3.1 given that N is normal, we can deduce that matrix of N is the 0 matrix, in other words $N = 0$.

exer:8A:14 14. (copy from part of exercise 13 (8A)) Since N is nilpotent, there exists a basis of V with respect to which the matrix of N is an upper triangular matrix with only 0's in the diagonal. Therefore, from theorem 8.3.4 (look at the proof of it, because only apply the theorem itself doesn't help), there exists an orthonormal basis of V with respect to which the matrix of N is an upper triangular matrix with only 0's in the diagonal.

exer:8A:15 15. Since $\text{null } N^{\dim V-1} \neq \text{null } N^{\dim V}$ so there exists $v \in V$ so $N^{\dim V}v = 0$ but $N^{\dim V-1}v \neq 0$. Therefore, according to exercise 5 (8A), $v, Tv, \ldots, T^{\dim V-1}v$ is linearly independent so it is basis of V. Note that $T^i v \in \text{null } T^{\dim V}$ so that means $V \subset \text{null } T^{\dim V}$ so $T^{\dim V} = 0$ so T is nilpotent.

Since $\text{null } N^{\dim V-1} \neq \text{null } N^{\dim V}$ so according to theorem 10.1.2 (8A), we obtain $\text{null } N^j \neq \text{null } N^{j+1}$ for all $0 \leq j \leq \dim V - 1$. Hence, from theorem 10.1.1 (8A), we obtain $\dim \text{null } N^j < \text{null } N^{j+1}$ for every $0 \leq j \leq \dim V - 1$, which leads to $\dim \text{null } N^i = i$ for every $1 \leq i \leq \dim V$.

exer:8A:16 16. For any nonnegative integer k, if $v \in \text{range } T^{k+1}$ then $v \in \text{range } T^k$ so $\text{range } T^{k+1} \subset \text{range } T^k$.

exer:8A:17 17. From exercise 16 (8A), we already know $\text{range } T^{m+k+1} \subset \text{range } T^{m+k}$. Now, if $v \in \text{range } T^{m+k}$ then $v = T^{m+k}u$. Since $T^m u \in \text{range } T^m = \text{range } T^{m+1}$ so there exists $w \in V$ so $T^m u = T^{m+1} w$ so $v = T^{m+k}u = T^{m+k+1}w$ so $v \in \text{range } T^{m+k+1}$. Thus, we obtain $\text{range } T^{m+k} \subset \text{range } T^{m+k+1}$. With this, we conclude $\text{range } T^{m+k} = \text{range } T^{m+k+1}$ for every nonnegative integer k, in other words $\text{range } T^k = \text{range } T^m$ for all $k > m$.

exer:8A:18 18. According to exercise 17 (8A), it suffices to prove $\text{range } T^n = \text{range } T^{n+1}$. Assume it is not true, then from exercises 16 and 17 (8A), we find

$$V = \text{range } T^0 \supsetneq \text{range } T^1 \supsetneq \cdots \supsetneq \text{range } T^{n+1}.$$

This follows $\dim \text{range } T^{n+1} < 0$, a contradiction. Thus, we must have $\text{range } T^n = \text{range } T^{n+1}$.

exer:8A:19 19. $\text{null } T^m = \text{null } T^{m+1}$ iff $\dim \text{null } T^m = \dim \text{null } T^{m+1}$ iff $\dim \text{range } T^m = \dim \text{range } T^{m+1}$ iff $\text{range } T^m = \text{range } T^{m+1}$.

exer:8A:20 20. From exercise 19 (8A), $\text{range } T^4 \neq \text{range } T^5$ follows $\text{null } T^4 \neq \text{null } T^5$ so from exercise 15 (8A), T is nilpotent.

exer:8A:21 21. $W \in \mathbf{C}^n$ so $T(z_1, \ldots, z_n) = (z_2, \ldots, z_n, 0)$ then $\text{null } T^i = \{(z_1, \ldots, z_i, 0, \ldots, 0) : z_i \in \mathbf{C}\}$.

10.3. 8B: Decomposition of an Operator

theo:8.21:8B **Theorem 10.3.1** (8.21, Description of operators on complex vector spaces) Suppose V is complex vector space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. Then:

(a) $V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$;

(b) each $G(\lambda_i, T)$ is invariant under T;

(c) each $(T - \lambda_i I)_{G(\lambda_i, T)}$ is nilpotent.

The idea for the proof of (a) is similar to of exercise 5 (5C).

Proof of (a). We induct on $\dim V$. From theorem 10.1.4 (8A) then $V = \text{null } (T - \lambda_i I)^n \oplus \text{range } (T - \lambda_i I)^n$ for all $1 \le i \le m$. Let $U = \text{range } (T - \lambda_i I)^n$ then $\dim U < \dim V$.

First, we show that for all $j \neq i$ then $G(\lambda_j, T) = G(\lambda_j, T|_U)$. It's obvious that $G(\lambda_j, T|_U) \subset G(\lambda_j, T)$. If $v \in G(\lambda_j, T)$ then $(T - \lambda_j I)^n v = 0$. With $V = \text{null } (T - \lambda_i I)^n \oplus \text{range } (T - \lambda_i I)^n$, we can write $v = u + (T - \lambda_i I)^n w$ where $u \in G(\lambda_i, T)$ so this follows $(T - \lambda_j I)^n u = 0 - (T - \lambda_j I)^n (T - \lambda_i I)^n w \in U$ so $(T - \lambda_j I)^n u \in U$. This follows there exists $s \in V$ so $(T - \lambda_j I)^n u = (T - \lambda_i I)^n s$. Note that $u \in G(\lambda_i, T)$ so that means $(T - \lambda_i I)^{2n} s = (T - \lambda_j I)^n (T - \lambda_i I)^n u = 0$ which leads to $(T - \lambda_i I)^n s = 0$ according to theorem 10.1.3 (8A) or $(T - \lambda_j I)^n u = 0$ or $u \in G(\lambda_j, T)$. Therefore,

$u \in G(\lambda_i, T) \cap G(\lambda_j, T)$ so from exercise 4 (8A) we find $u = 0$. Thus, $v = (T - \lambda_i I)^n w \in U$. We conclude that any generalized eigenvector v of T corresponding to $\lambda_j \neq \lambda_i$ is in U, which means $v \in G(\lambda_j, T|_U)$. With this, we deduce, $G(\lambda_j, T) \subset G(\lambda_j, T|_U)$ so $G(\lambda_j, T) = G(\lambda_j, T|_U)$ for all $j \neq i$.

Second, note that λ_j ($j \neq i$) are the only eigenvalues of operator $T|_U$ over U. Hence, since $\dim U < \dim V$ so from inductive hypothesis, we obtain

$$\text{range } (T - \lambda_i I)^n = U = \bigoplus_{j \neq i} G(\lambda_j, T|_U) = \bigoplus_{j \neq i} G(\lambda_j, T).$$

Thus, $V = G(\lambda_i, T) \oplus \text{range } (T - \lambda_i I)^n = \bigoplus_{j=1}^m G(\lambda_j, T)$. □

theo:8.26:8B **Theorem 10.3.2** (8.26) Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Then the sum of multiplicities of all eigenvalues of T equals $\dim V$.

theo:8.29:8B **Theorem 10.3.3** (8.29) Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T, with multiplicities $d_1, \ldots, d_m$. Then there is a basis of V with respect to which T has a block diagonal matrix of the form

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

where each A_j is a d_j-by-d_j upper triangular matrix of the form

$$A_j = \begin{pmatrix} \lambda_j & & \star \\ & \ddots & \\ 0 & & \lambda_j \end{pmatrix}.$$

theo:8.31:8B **Theorem 10.3.4** (8.31) Suppose $N \in \mathcal{L}(V)$ is nipotent. Then $I + N$ has kth root.

theo:8.33:8B **Theorem 10.3.5** (8.33) Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ is invertible. Then T has a kth root.

10.4. Exercises 8B

exer:8B:1 1. Since 0 is the only eigenvalue of N and that V is a complex vector space, from theorem 7.3.2 (5B) and 7.3.4 (5B), we follow there exists a basis of V with respect to which the matrix of N is an upper-triangular matrix with only 0's on the diagonal. From exercise 12 (8A), we follow N is nilputent.

exer:8B:2 2. The idea is choose operator T so T has 0 as the only eigenvalue and that there doesn't exists any basis of V with respect to which T has upper-triangular matrix. Hence:

- First pick null $T = \text{span}\,(v_1)$.
- Choose $U = \text{range}\, T$ so $T|_U \in \mathcal{L}(U)$ does not have any eigenvalue. Let $\dim U = 2$ then say $Tv_2 = v_3, Tv_3 = -v_2$, we can find that $T|_U$ does not have any eigenvalue.
- Confirm that T is not nilpotent. Now we have $V = \text{span}\,(v_1, v_2, v_3)$ so it suffices to show $T^3 \neq 0$. Indeed, we have

$$T^3(av_1 + bv_2 + cv_3) = T^2(bv_3 - cv_2) = T(-bv_2 - cv_3) = -bv_3 + cv_2.$$

exer:8B:3 3. If λ is eigenvalue of T and $v \in E(\lambda, T)$ then since S is invertible, there exists $u \in V$ so $Su = v$. This follows $S^{-1}TSu = S^{-1}Tv = \lambda S^{-1}v = \lambda u$ so λ is also eigenvalue of T. Similarly, if λ is eigenvalue of $S^{-1}TS$ and $v \in E(\lambda, S^{-1}TS)$ then $TSv = \lambda Sv$ so λ is also eigenvalue of T. Thus, T and $S^{-1}TS$ have same eigenvalues.

Now it suffices to show that $\dim G(\lambda, T) = \dim G(\lambda, S^{-1}TS)$ for every eigenvalue of λ. Note that from exercise 5 (5B) we have $(S^{-1}TS - \lambda I)^n = S^{-1}(T - \lambda I)^n S$. Let $v_1, \ldots, v_m$ be basis of $G(\lambda, S^{-1}TS)$ then $(S^{-1}TS - \lambda I)^n v_i = S^{-1}(T - \lambda I)^n Sv_i = 0$ so $(T - \lambda I)^n Sv_i = 0$. This follows $Sv_i \in G(\lambda, T)$ and also observe that $Sv_1, \ldots, Sv_m$ is linearly independent so that means $\dim G(\lambda, S^{-1}TS) \leq \dim G(\lambda, T)$. Similarly, if $Sv_1, \ldots, Sv_m$ is basis of $\dim G(\lambda, T)$ then $v_1, \ldots, v_m \in G(\lambda, S^{-1}TS)$ and $v_1, \ldots, v_m$ is linearly independeny. With this, we obtain $\dim G(\lambda, T) \leq \dim G(\lambda, S^{-1}TS)$. Hence, we obtain $\dim G(\lambda, T) = \dim G(\lambda, S^{-1}TS)$.

exer:8B:4 4. If null $T^{n-2} \neq$ null T^{n-1} then from theorem 10.1.2 (8A), we find null $T^i \neq$ null T^{i+1} for all $0 \leq i \leq n-2$. This follows $\dim G(0, T) \geq n - 1$, i.e. the multiplicity of 0 of T is at least $n - 1$. Hence, from theorem 10.3.2 (8B), T has at most 2 eigenvalues.

exer:8B:5 5. If every generalized eigenvector of T is an eigenvector of T then obviously V is has basis consisting of eigenvectors of T. For the other direction, if V has basis consisting of eigenvectors of T then that means $\sum_{i=1}^m \dim E(\lambda_i, T) = \sum_{i=1}^m \dim G(\lambda_i, T)$ and the equality holds when $E(\lambda_i, T) = G(\lambda_i, T)$ for every $1 \leq i \leq m$, i.e. every generalized eigenvector of T is an eigenvector of T.

exer:8B:6 6. We have $N^5 = 0$ so N is nilpotent. According to proof of theorem 10.3.4 (8B), let $R = I + a_1N + a_2N^2 + a_3N^3 + a_4N^4$ be the square root of $N + I$ then

$$R^2 = I + 2a_1N + (a_1^2 + 2a_2)N^2 + (2a_3 + 2a_1a_2)N^3 + (2a_4 + 2a_3a_1 + a_2^2)N^4.$$

We find $a_1 = 1/2, a_2 = -1/8, a_3 = 1/16, a_4 = -5/128$.

exer:8B:7 7. First is show that $I + N$ where N is nilpotent has a cube root then follow similarly to proof of theorem 10.3.5 (8B).

exer:8B:8

8. If $3, 8$ are the only eigenvalues of T then null $T^{n-2} =$ null $T^0 = \{0\}$ so according to exercise 19 (8A), range $T^{n-2} =$ range $T^0 = V$.

   If T has at least 3 eigenvalues then from exercise 4 (8B), we obtain null $T^{n-2} =$ null $T^{n-1} =$ null T^n so from exercise 19 (8A) again, we have range $T^{n-2} =$ range $T^{n-1} =$ range T^n. Thus, from theorem 10.1.4 (8A), we find $V =$ null $T^n \oplus$range $T^n =$ null $T^{n-2} \oplus$range T^{n-2}.

exer:8B:9

9. Not hard.

exer:8B:10

10. Since T is an operator on a complex vector space, from theorem 10.3.3 (8B), there exists a basis of V with respect to which T has a block diagonal matrix.

    Let D_i be a diagonalizable d_j-by-d_j matrix with its diagonal same as matrix A_j. Then let $D \in \mathcal{L}(V)$ so matrix of D is
    $$\begin{pmatrix} D_1 & & 0 \\ & \ddots & \\ 0 & & D_m \end{pmatrix}.$$
    Let $N_i = A - D_i = \lambda_i I_i$ where I_i is d_i-by-d_i identity matrix. Let $N \in \mathcal{L}(V)$ so matrix of N is
    $$\begin{pmatrix} N_1 & & 0 \\ & \ddots & \\ 0 & & N_m \end{pmatrix} = \begin{pmatrix} \lambda_1 I_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m I_m \end{pmatrix}$$
    From here, we obtain N is nilpotent, $T = N + D$ and from exercise 9 (8B), we have matrix of ND is
    $$\begin{pmatrix} \lambda_1 I_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m I_m \end{pmatrix} \begin{pmatrix} D_1 & & 0 \\ & \ddots & \\ 0 & & D_m \end{pmatrix} = \begin{pmatrix} \lambda_1 D_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m D_m \end{pmatrix},$$
    which is also equal to matrix of DN. Thus, $ND = DN$.

exer:8B:11

11. First we show that number of times that 0 appears on the diagonal of the matrix of T is at most dim null T^n. For a basis of V with respect to which T has an upper-triangular matrix, let $v_{a_1}, \ldots, v_{a_m}$ be subset of that basis so the i-th 0 on the diagonal (from top left to bottom right) is also in a_i-th column of matrix of T.

    We prove inductively on i that there exist $u_i \in$ span $(v_1, \ldots, v_{a_i})$ so $Tu_1 = 0$ and $Tu_i \in$ span $(u_1, \ldots, u_{i-1})$ for $i \geq 2$. For $i = 1$, let $u_1 = \sum_{i=1}^{a_1-1} \alpha_i v_i + v_{a_1}$. Since $Tv_{a_1} \in$ span $(v_1, \ldots, v_{a_1-1})$ so we can write
    $$T(u_1) = T\left(\sum_{i=1}^{a_1-1} \alpha_i v_i\right) + \sum_{i=1}^{a_1-1} \beta_i v_i = \sum_{i=1}^{a_1-1} (\alpha_i T v_i + \beta_i v_i).$$
    I suffices to choose α_i so $T(u_1) = 0$. Indeed, observe that the first $a_1 - 1$ numbers on the diagonal (from top left) are nonzero, i.e. for every $1 \leq i \leq a_1 - 1$ then $Tv_i = \sum_{j=1}^{i} c_{j,i} v_j$

with $c_{j,i} \neq 0$ for all $1 \le j \le i$. Hence,

$$T(u_1) = \sum_{i=1}^{a_1-1} \left(\sum_{j=i}^{a_1-1} \alpha_j c_{i,j} + \beta_i \right) v_i.$$

Therefore, since $c_{i,j} \neq 0$ for all $1 \le i \le a_1 - 1, 1 \le j \le i$, we can choose α_{a_1-1} so $\alpha_{a_1-1} c_{a_1-1,a_1-1} + \beta_{a_1-1} = 0$, then we can choose α_{a_1-2} so $\sum_{j=a_1-2}^{a_1-1} \alpha_j c_{a_1-2,j} + \beta_{a_i-2} = 0$. Continuing this, we guarantee to find u_1 so $Tu_1 = 0$.

If the statement is true for all numbers less than i, consider $u_i = v_{a_i} + \sum_{j=1}^{a_i-1} \alpha_j v_j$. We show that we can find α_j so $Tu_i \in \text{span}\ (u_1, \ldots, u_{i-1})$. Indeed, we have

$$Tu_i = T\left(\sum_{j=1}^{a_{i-1}} \alpha_j v_j \right) + T\left(\sum_{j=a_{i-1}+1}^{a_i-1} \alpha_j v_j \right) + Tv_{a_i}.$$

Observe that on the diagonal, there is no 0 between a_{i-1}-th number and a_i-th number, so with similar approach to case $i = 1$, we can choose $\alpha_{a_{i-1}+1}, \ldots, \alpha_{a_i-1}$ so scalar multiplication of v_j $(a_{i-1} + 1 \le j \le a_i - 1)$ in representation of Tu_i is 0. With this and note that $u_{i-1} \in \text{span}\ (v_1, \ldots, v_{a_{i-1}})$, we obtain:

$$\begin{aligned} Tu_i &= T\left(\sum_{j=1}^{a_{i-1}} \alpha_j v_j \right) + \sum_{i=1}^{a_{i-1}} \beta_i v_i, \\ &= T\left(\sum_{j=1}^{a_{i-1}} \alpha_j v_j \right) + \sum_{i=1}^{a_{i-1}-1} k_i v_i + \lambda_{i-1} u_{i-1}, \\ &= T\left(\sum_{j=1}^{a_{i-2}} \alpha_j v_j \right) + T\left(\sum_{j=a_{i-2}+1}^{a_{i-1}} \alpha_j v_j \right) + \sum_{i=1}^{a_{i-1}-1} k_i v_i + \lambda_{i-1} u_{i-1}. \end{aligned}$$

where $k_i \in \mathbf{F}$. Similarly, we can choose $\alpha_{a_{i-2}+1}, \ldots, \alpha_{a_{i-1}}$ so scalar multiplications of v_j $(a_{i-2} + 1 \le j \le a_{i-1} - 1$ in Tu_i are 0. With this, we obtain:

$$Tu_i = T\left(\sum_{j=1}^{a_{i-2}} \alpha_j v_j \right) + \sum_{i=1}^{\alpha_{i-2}-1} \ell_i v_i + \lambda_{i-2} u_{i-2} + \lambda_{i-1} u_{i-1}.$$

By continuing this method, we can deduce $Tu_i \in \text{span}\ (u_1, \ldots, u_{i-1})$.

From the above, since $Tu_1 = 0$ and $Tu_i \in \text{span}\ (u_1, \ldots, u_{i-1})$, we obtain $u_1, \ldots, u_m \in$ null T^n. On the other hand, we also find that $u_1, \ldots, u_m$ is linearly independent so that means $m \le \dim \text{null} T^n$.

Back to our original question, for λ_i as an eigenvalue of T, $\mathcal{M}(T)$ is an upper-triangular matrix with respect to some basis of V then $\mathcal{M}(T - \lambda_i I)$ is also an upper-triangular matrix. Furthermore, number of times λ_i appears on the diagonal of $\mathcal{M}(T)$ (denote this as S_i) equals to number of times 0 appears on the diagonal of $\mathcal{M}(T-\lambda_i I)$, which is at most dim null $(T-\lambda_i I)^n$ times according to what we've proven above. We write $S_i \le$ dim null $(T-\lambda_i I)^n$. On the other hand, $\sum_i S_i = \sum_i \dim \text{null } (T-\lambda_i I)^n = n$ so that means $S_i = \dim G(\lambda_i, T)$ for all i. We're done.

10.5. 8C: Characteristic and Minimal Polynomials

Definition 10.5.1 (8.34, characteristic polynomial)**.** Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_m$ denote the distinct eigenvalues of T, with multiplicities $d_1, \dots, d_m$. The polynomial $(z-\lambda 1)^{d_1} \cdots (z-\lambda_m)^{d_m}$ is called the characteristic polynomial of T.

theo:8.37:8C:cayley_hamilton **Theorem 10.5.2** (Cayley-Hamilton Theorem) Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let q denote the characteristic polynomial of T. Then $q(T) = 0$.

Definition 10.5.3 (8.43, minimal polynomial)**.** Suppose $T \in \mathcal{L}(V)$. Then the minimal popynomial of T is the unique monic polynomial p of smallest degree such that $p(T) = 0$.

theo:8.46:8C **Theorem 10.5.4** (8.46) Suppose $T \in \mathcal{L}(V)$ and $q \in \mathcal{P}(\mathbf{F})$. Then $q(T) = 0$ if and only if q is a polynomial multiple of the minimal polynomial of T.

With $\mathbf{F} = \mathbf{C}$, theorem 10.5.4 and Cayley-Hamilton Theorem 10.5.2 follows that the charactersitic polynomial of T is a polynomial multiple of the minimal polynomial of T.

theo:8.49:8C **Theorem 10.5.5** (8.49) Let $T \in \mathcal{L}(V)$. Then the zeros of the minimal polynomial of T are precisely the eigenvalues of T.

Not from the book:

theo:3:8C **Theorem 10.5.6** Let $p(z) = (z-\lambda_1)^{d_1} \cdots (z-\lambda_m)^{d_m}$ be the **minimal** polynomial of T then $G(\lambda_i, T) = \text{null } (T-\lambda_i I)^{d_i}$ for all $1 \le i \le m$. In particular, d_i is the smallest number so null $(T-\lambda_i I)^{d_i} =$ null $(T-\lambda_i)^n$, or we can say that $(z-\lambda_i)^{d_i}$ is the minimal polynomial of $T|_{G(\lambda_i,T)}$.

The idea for the proof is similar to proof of exercise 12 (8C).

How to find characteristic/minimal polynomial of an operator.

Question 10.5.7. What can you tell about an operator from looking at its minimal/characteristic polynomial?

Good summary question

10.6. Exercises 8C

exer:8C:1 1. If $3, 5, 8$ are the only eigenvalues of T then each of them has multiplicity of at most 2. This follows characteristic polynomial of T is $(z-3)^{d_1}(z-5)^{d_2}(z-8)^{d_3}$ with $d_i \le 2$. Hence, $(T-3I)^2(T-5I)^2(T-8I)^2 = 0$.

exer:8C:2 2. If $5, 6$ are the only eigenvalues of T then each of them has multiplicity of at most $n-1$. This follows $(T-5I)^{n-1}(T-6I)^{n-1} = 0$.

exer:8C:3 3. Let $T(1,0,0,0) = 7(1,0,0,0), T(0,1,0,0) = 7(0,1,0,0)$ and $T(0,0,1,0) = 8(0,0,1,0)$, $T(0,0,0,1) = 8(0,0,0,1)$.

exer:8C:4 4. If the minimal polynomial is $(z-1)(z-5)^2$, this could mean that $G(5,T) = \text{null } (T-5I)^2$. Choose $T(a,b,c,d) = (a, 5b, 5c, c+5d)$ then $T^2(a,b,c,d) = (a, 25b, 25c, 10c+25d)$. Hence, we find $G(1,T) = \text{span }((1,0,0,0))$ and $G(5,T) = \text{span }((0,1,0,0),(0,0,1,0),(0,0,0,1))$, in particular $(T-5I)(0,0,1,0) \neq 0$ but $(T-5I)(0,0,1,0) = 0$; $(T-5I)(0,1,0,0) = (T-5I)(0,0,0,1) = 0$ so indeed $G(5,T) = \text{null } (T-5I)^2$. The characteristic polynomial equals $(z-1)(z-5)^3$.

   Note that $(z-1)(z-5)$ is not minimal polynomial because $(T-I)(T-5I)^2(0,0,1,0) \neq 0$ and $(T-I)(T-5I)^2 = 0$ because $G(5,T) = \text{null } (T-5I)^2$. Hence $(z-1)(z-5)^2$ is the minimal polynomial.

exer:8C:5 5. This means $G(1,T) = \text{null } (T-I)^2 \neq \text{null } (T-I)$. Hence, choose $T(a,b,c,d) = (0, 3b, c, c+d)$.

exer:8C:6 6. This means $G(1,T) = \text{null } (T-I)$. Choose $T(a,b,c,d) = (0,3b,c,d)$.

exer:8C:7 7. From $P^2 = P$ we follow $0, 1$ are the only eigenvalues of P. Since $P^2 = P$ so $\dim G(0,T) = \dim \text{null } P$ so the characteristic polynomial of P is $z^{\dim \text{null } P}(z-1)^{\dim V - \dim \text{null } P}$. From $P^2 = P$ we also find that $V = \text{range } P \oplus \text{null } P$ according to exercise 4 (5B). Hence $\dim \text{null } P + \dim P = \dim V$ so the characteristic polynomial of P is $z^{\dim \text{null } P}(z-1)^{\dim \text{range } P}$.

exer:8C:8 8. T is invertible iff 0 is not an eigenvalue of T iff 0 iso not a zero of the minimal polynomial of T iff the constant term in the minimal polynomial of T is nonzero.

exer:8C:9 9. If $p(z) = 4 + 5z - 6z^2 - 7z^3 + 2z^4 + z^5$ is minimal polynomial of T then $q(z) = \frac{1}{4}z^5 p(z^{-1})$ the is minimal polynomial of T^{-1}. Indeed, because $p(T)v = 0$ for all $v \in V$ so $q(T^{-1})v = T^{-5}p(T)v = 0$ for all $v \in V$. Hence, from theorem 10.5.4 (8C), $q(z)$ is polynomial multiple of the minimal polynomial of T^{-1}.

   If the minimal polynomial of T^{-1} has degree less than degree of $q(z)$ then with similar construction as above, we can show that there exists a minimal polynomial of T whose degree is less than degree of $q(z)$, which is 5, a contradiction. Thus, we obtain $q(z)$ is the minimal polynomial of T^{-1}.

exer:8C:10 10. Since T is invertible so the constant term in the characteristc polynomial of T is nonzero, i.e. $p(0) \neq 0$. Let $\lambda_1, \ldots, \lambda_m$ be eigenvalues of T with corresponding multiplicities $d_1, \ldots, d_m$. Hence, $p(z) = (z-\lambda_1)^{d_1} \cdots (z-\lambda_m)^{d_m}$. From exercise 3 (8A), we have

$G(\lambda_i, T) = G(\frac{1}{\lambda_i}, T^{-1})$ so characteristic polynomial of T^{-1} is

$$\begin{aligned} q(z) &= (z-\lambda_1^{-1})^{d_1}\cdots(z-\lambda_m^{-1})^{d_m}, \\ &= \lambda_1^{-d_1}\cdots\lambda_m^{-d_m}\cdot z^{\dim V}\cdot(\lambda_1 - z^{-1})^{d_1}\cdots(\lambda_m - z^{-1})^{d_m}, \\ &= \frac{(-1)^{\dim V}}{p(0)}\cdot z^{\dim V}(\lambda_1 - z^{-1})^{d_1}\cdots(\lambda_m - z^{-1})^{d_m}, \\ &= \frac{1}{p(0)} z^{\dim V} p\left(\frac{1}{z}\right). \end{aligned}$$

exer:8C:11 11. Since T is invertible so the minimal polynomial of T $p(z) = a_0 + a_1 z + \ldots + a_m z^m$ has $a_0 \neq 0$ according to exercise 8 (8C). Therefore $T^{-1}p(T) = 0$ or $a_0 T^{-1} = -a_1 I - a_2 T - \ldots - a_m T^{m-1}$.

exer:8C:12 12. According to exercise 5 (8B), V has basis consisting of eigenvectors of T if and only if every generalized eigenvector of T is an eigenvector of T.

Let $p(z) = (z-\lambda_1)^{d_1}\cdots(z-\lambda_m)^{d_m}$ be the minimal polynomial of T. Since $p(T)v = 0$ for all $v \in V$ so that means $\prod_{j\neq i}(T-\lambda_j I)^{d_j} v \in \text{null } (T-\lambda_i I)^n$ for all $v \in V$.

Next, we show that $p(z)$ has no repeated zero iff all generalized eigenvector of T is an eigenvalue of T. Indeed, if $p(z)$ has a repeated zero λ_i, i.e. $d_i \geq 2$ then there exists $v \in V$ so $v \notin \text{null } (T-\lambda_i I)$ but $v \in \text{null } (T-\lambda_i I)^{d_i}$, otherwise $\prod_{j\neq i}(T-\lambda_j I)^{d_j} v \in \text{null } (T-\lambda_i I)$ for all $v \in V$ which leads to $q(T) = (T-\lambda_i I)\prod_{j\neq i}(T-\lambda_j I)^{d_j} = 0$, where $q(z)$ is a polynomial with degree less than $p(z)$, a contradiction. Conversely, if there exists a generalized eigenvector v of T corresponding to eigenvalues λ_i but v is not an eigenvector of T corresponding to λ_i, i.e. $(T-\lambda_i I)v \neq 0$ but $v \in G(\lambda_i, T)$. Note that for any $1 \leq j \leq m$ then $T-\lambda_j I$ is invariant under $G(\lambda_i, T)$. Combining with the fact that $G(\alpha, T)\cap G(\beta, T) = \{0\}$ for $\alpha \neq \beta$ from exercise 4 (8A), we follow that $\prod_{j\neq i}(T-\lambda_j I)^{d_j}(T-\lambda_i I)v \neq 0$ and $\prod_{j\neq i}(T-\lambda_j I)^{d_j}(T-\lambda_i I)v \in G(\lambda_i, T)$. This follows $p(z)$ has a repeated zero λ_i.

From these two claims, we are done.

exer:8C:13 13. If $\mathbf{F} = \mathbf{C}$ then according Complex Spectral Theorem 9.3.1 (7B), T has a diagonal matrix with respect to some orthonormal basis of V which means V has a basis consisting of eigenvectors of T. Hence, according to previous exercise, the minimal polynomial of T has no repeated zero.

exer:8C:14 14. Since S is an isometry so according to theorem 9.5.4 (7C), there exists a basis of V consisting of eigenvalues of S whose corresponding eigenvectors all have absolute value 1. On the other hand, the constant term in the characteristic polynomial of S is $p(0) = (-1)^n \lambda_1^{d_1}\cdots\lambda_m^{d_m}$ so $|p(0)| = 1$.

exer:8C:15 15. (a) We know that there exists a monic polynomial $p(z)$ (such as minimal polynomial of T) so $p(T)v = 0$. Hence, it suffices the uniqueness of such polynomial of smallest degree. Assume the contrary, there exists two polynomial p, q with same and smallest degree so

$p(T)v = q(T)v = 0$. We follow $(p-q)(T)v = 0$ but $p - q \neq 0$ and $p - q$ has degree less than of p, q, which is a contradiction. Thus, uniqueness is proven.

(b) Let q be the minimal polynomial of T. We follow hat $\deg p \leq \deg q$. Hence, there exists polynomials r, s so $\deg s < \deg p$ and $q = pr + s$. Since $p(T)v = q(T)v = 0$ so $s(T)v = 0$ but $\deg s < \deg p$ so $s = 0$. This follows p divides the minimal polynomial of T.

exer:8C:16 16. Let $p(z) = a_0 + a_1 z + \ldots + a_{m-1} z^{m-1} + a_m$ be the minimal polynomial of T and $q(z) = \overline{a_0} + \overline{a_1} z + \ldots + \overline{a_{m-1}} z^{m-1} + a_m$. Then we find $q(T^*) = (p(T))^*$. Note that since $p(T)v = 0$ for all $v \in V$ so $0 = \langle p(T)q(T^*)v, v\rangle = \langle q(T^*)v, (p(T))^* v\rangle = \|q(T^*)v\|^2$ for all $v \in V$ so $q(T^*)v = 0$. Therefore, from theorem 10.5.4 (8C), q is the polynomial multiple of the minimal polynomial h of T^*.

Now it suffices to show that $\deg q = \deg h$. Assume the contrary $\deg h < \deg q$ then with similar approach, we can construct a polynomial g from h so $g(T) = 0$ and $\deg g = \deg h < \deg q = \deg p$, a contradiction. Thus, we must have $\deg g = \deg h$ so $q(z)$ is the minimal polynomial of T^*.

exer:8C:17 17. Let $q(z) = (z - \lambda_1)^{k_1} \cdots (z - \lambda_m)^{k_m}$ be the minimal polynomial of T with $k_1 + \ldots + k_m = n$ and $p(z) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_m)^{d_m}$ be the characteristic polynomial of T. According to theorem 10.5.4 (8C), we follow $d_i \geq k_i$ for all $1 \leq i \leq m$. On the other hand, $\sum_{i=1}^m k_i = \sum_{i=1}^m d_i = n$ so that means $d_i = k_i$ for all $1 \leq i \leq m$, i.e. the characteristic polynomial of T equals the minimal polynomial of T.

exer:8C:18 18. Call such operator T and let $e_1, \ldots, e_n$ be the standard basis of $\mathbf{C}^n$ with e_i has 1 in the i-th coordinate and 0's in the remaining coordinates. We can easily find that for $1 \leq i \leq n-1, 1 \leq k \leq n$ then $T^k e_i = e_{i+k}$ if $i + k \leq n$.

We have $Te_n = -(a_0, a_1, \ldots, a_{n-1})$ so

$$T^2 e_n = -T(a_0, \ldots, a_{n-2}, 0) - a_{n-1} T e_n = -(0, a_0, \ldots, a_{n-2}) - a_{n-1} T e_n.$$

Hence,

$$\begin{aligned} T^3 e_n &= -T(0, a_0, \ldots, a_{n-3}, 0) - a_{n-2} T e_n - a_{n-1} T^2 e_n, \\ &= -(0, 0, a_0, \ldots, a_{n-3}) - a_{n-2} T e_n - a_{n-1} T^2 e_n. \end{aligned}$$

Continuing doing this, we will obtain

$$T^i e_n = -(0, 0, \ldots, 0, a_0, \ldots, a_{n-i}) - a_{n-i+1} T e_n - \ldots - a_{n-1} T^{i-1} e_n.$$

For $i = n$ then $T^n e_n = -a_0 e_n - a_1 T e_n - \ldots - a_{n-1} T^{n-1} e_n$ or $(T^n + a_{n-1} T^{n-1} + \ldots + a_1 T + a_0 I) e_n = p(T) e_n = 0$.

For $1 \leq i < n$, notice that $T^k e_i = e_{i+k}$ for $0 \leq k \leq n - i$. Hence, we have

$$\begin{aligned} p(T) e_i &= (T^n + a_{n-1} T^{n-1} + \ldots + a_{n-i+1} T^{n-i+1}) e_i + a_{n-i} e_n + \ldots + a_1 e_{i+1} + a_0 e_i, \\ &= (T^i + a_{n-1} T^{i-1} + \ldots + a_{n-i+1} T)(T^{n-i} e_i) + (0, \ldots, 0, a_0, \ldots, a_{n-i-1}), \\ &= (T^i + a_{n-1} T^{i-1} + \ldots + a_{n-i+1} T) e_n + (0, \ldots, 0, a_0, \ldots, a_{n-i}), \\ &= 0. \end{aligned}$$

Thus, $p(T)v = 0$ for all $v \in V$. Since $\deg p = n$ so that means $p(z)$ is both the minimal polynomial and the characteristic polynomial of T.

exer:8C:19 19. According to exercise 11 (8B), the number of times that λ appears on the diagonal of the matrix of T equals the multiplicity of λ as an eigenvalue of T. Combining with the definition of characteristic polynomial for complex vector space, we obtain the characteristic polynomial of T is $(z-\lambda_1)\cdots(z-\lambda_n)$.

exer:8C:20 20. Since we are talking about complex vector space, according to theorem 7.3.2 (5B), $T|_{V_j}$ has an upper-triangular matrix $\mathcal{M}(T|_{V_j})$ with respect to some basis of V_j.

Therefore, if we combines all the bases of V_j we obtain a basis of V since $V = V_1 \oplus \cdots \oplus V_m$ which will give an block diagonal matrix for T as follow:

$$\mathcal{M}(T) = \begin{pmatrix} \mathcal{M}(T|_{V_1}) & & \\ & \ddots & \\ 0 & & \mathcal{M}(T|_{V_m}) \end{pmatrix}.$$

Notice that $\mathcal{M}(T)$ is an upper-triangular matrix since $\mathcal{M}(T|_{V_j})$ are upper-triangular matrices. Thus, by applying exercise 19 (8C) to $\mathcal{M}(T), \mathcal{M}(T|_{V_j})$ we can conclude that $p_1 \cdots p_m$ is the characteristic polynomial of T.

10.7. 8D: Jordan Form

theo:8.55:8D **Theorem 10.7.1** (8.55, basis correspoding to a nilpotent operator) Suppose $N \in \mathcal{L}(V)$ is nilpotent. There exists vectors $v_1, \dots, v_n \in V$ and nonnegative $m_1, \dots, m_n$ such that

(a) $N^{m_1}v_1, \dots, Nv_1, v_1, \dots, N^{m_n}v_n, \dots, Nv_n, v_n$ is a basis of V.

(b) $N^{m_1+1}v_1 = \dots = N^{m_n+1}v_n = 0$.

Furthermore, according to exercise 6 (8D), $n = \dim \operatorname{null} N$ and $N^{m_1}v_1, \dots, N^{m_n}v_n$ is basis of null N.

theo:8.60:8D **Theorem 10.7.2** (8.60, Jordan Form) Suppose V is a complex vector space. If $T \in \mathcal{L}(V)$, then there is a basis of V that is a Jordan basis for T.

In particular, there is a Jordan basis of V for T so

$$\mathcal{M}(T) = \begin{pmatrix} \mathcal{M}(T|_{G(\lambda_1,T)}) & & 0 \\ & \ddots & \\ 0 & & \mathcal{M}(T|_{G(\lambda_m,T)}) \end{pmatrix}.$$

In here, $\mathcal{M}(T|_{G(\lambda_i,T)})$ is Jordan matrix of $T|_{G(\lambda_i,T)}$ on $G(\lambda_i, T)$, i.e.

$$\mathcal{M}(T|_{G(\lambda_i,T)}) = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix},$$

where A_j is an upper-triangular matrix of the form:

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

Indeed, since $(T - \lambda_i I)|_{G(\lambda_i,T)}$ is nilpotent so from theorem 10.7.1, there exists such basis to form Jordan matrix for $T|_{G(\lambda_i,T)}$.

theo:4:8D **Theorem 10.7.3** The minimal polynomial of T is $\prod_{i=1}^{m}(z-\lambda_i)^{m_i+1}$ where m_i is the length of the longest consecutive string of 1's that appears on the line directly above the diagonal in the Jordan matrix of $\mathcal{M}(T|_{G(\lambda_i,T)})$.

Proof. From theorem 10.5.6 (8C), minimal polynomial of T equals product of minimal polynomials of $T|_{G(\lambda_i,T)}$ for all eigenvalues λ_i of T.

From exercise 3 (8D), minimal polynomial of the nilpotent operator $(T-\lambda_i I)|_{G(\lambda_i,T)}$ is z^{m_i+1} where m_i is the length of the longest consecutive string of 1's that appears on the line directly above the diagonal in the Jordan matrix of $\mathcal{M}(T|_{G(\lambda_i,T)})$. We follow $(T-\lambda_i I)^{m_i+1} = 0$ but $(T-\lambda_i i)^{m_i} \neq 0$, which leads to $(z-\lambda_i)^{m_i+1}$ as the minimal polynomial of $T|_{G(\lambda_i,T)}$.

Thus, the minimal polynomial of T is $\prod_{i=1}^{m}(z-\lambda_i)^{m_i+1}$ where m_i is the length of the longest consecutive string of 1's that appears on the line directly above the diagonal in the Jordan matrix of $\mathcal{M}(T|_{G(\lambda_i,T)})$. ∎

example_jordan_form **Example 10.7.4** we can find the minimal polynomial of T from its Jordan matrix as shown:

$$\mathcal{M}(T) = \begin{array}{c} \\ N^2v_1 \\ Nv_1 \\ v_1 \\ Nv_2 \\ v_2 \\ Nv_3 \\ v_3 \end{array} \begin{array}{c} \begin{array}{ccccccc} N^2v_1 & Nv_1 & v_1 & Nv_2 & v_2 & Nv_3 & v_3 \end{array} \\ \begin{pmatrix} \lambda_1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 \end{pmatrix} \end{array}$$

So $N^2v_1, Nv_1, v_1, Nv_2, v_2 \in G(\lambda_1, T)$ and $Nv_3, v_3 \in G(\lambda_2, T)$. In the top left submatrix of $\mathcal{M}(T)$ that only has λ_1 on the diagonal, the longest length of a consecutive strings of 1's is 2, which follows $(z-\lambda_1)^3$ is the minimal polynomial of $T|_{G(\lambda_1,T)}$ according to theorem 10.7.3. Similarly, $(z-\lambda_2)^2$ is the minimal polynomial of $T|_{G(\lambda_2,T)}$. Thus, the minimal polynomial of T is $(z-\lambda_1)^3(z-\lambda_2)^2$.

10.8. Exercises 8D

exer:8D:1

1. Since N is nilpotent so $N^4 = 0$ and since $N^3 \neq 0$, we find z^4 is both the minimal polynomial and the characteristic polynomial of T.

exer:8D:2

2. In example 8.54, with the basis $N^2v_1, Nv_1, v_1, Nv_2, v_2, v_3$ given that $N^3v_1 = N^2v_2 = Nv_3 = 0$, we follows that $N^3 = 0$ and $N^2 \neq 0$. Hence, N^3 is the minimal polynomial of T and N^6 is the characteristic polynomial of T.

exer:8D:3

3. We know so far according to theorem 10.7.1 (8D) and 10.7.2 (8D), Jordan basis of V is a combination of linear independent lists of the form $N^m v, N^{m-1}v, \dots, Nv, v$. Each of this list will produce a consecutive string of 1's with length m that appears on the line directly above the diagonal in the Jordan matrix of T. Therefore, the longest consecutive string of 1's corresponds to the list $N^m v, N^{m-1}v, \dots, Nv, v$ with maximal m. We follow $N^{m+1} = 0$ and $N^m \neq 0$ (because $N^m v$ is in the Jordan basis of V). Therefore, z^{m+1} is the minimal polynomial of T. Note that m here is the length of the longest consecutive string of 1's.

exer:8D:4

4. If the Jordan basis $v_1, \dots, v_n$ is reversed then the columns of $N^j v_i$ are reversed, i.e. if it is $(a_1, a_2, \dots, a_n)^T$ then its reverse is $(a_n, \dots, a_1)^T$. All of the columns are then arranged in reversed order. Thus, with this, we conclude that if we rotate 180° the Jordan matrix of T, we will obtain a matrix with respect to the reversed order of Jordan basis of V.

exer:8D:5

5. In the matrix of T^2 with respect to Jordan basis of V for T, all the eigenvalues on the diagonal are squared and all the 1's that is in the same column with λ changes to 2λ. All the 0's in row of $N^k v_i$ and column $N^{k-2}v_i$ changes to 1. Below is an example:

$$\mathcal{M}(T^2) = \begin{array}{c} \\ N^2v_1 \\ Nv_1 \\ v_1 \\ Nv_2 \\ v_2 \\ Nv_3 \\ v_3 \end{array} \begin{array}{c} \begin{array}{ccccccc} N^2v_1 & Nv_1 & v_1 & Nv_2 & v_2 & Nv_3 & v_3 \end{array} \\ \begin{pmatrix} \lambda_1^2 & 2\lambda_1 & 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1^2 & 2\lambda_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2^2 & 2\lambda_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_3^2 & 2\lambda_3 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_3^2 \end{pmatrix} \end{array}$$

exer:8D:6

6. Since we already know that $N^{m_1}v_1, \dots, N^{m_n}v_n$ is linearly independent list in null N, it suffices to show that this list spans null N. Indeed, consider $v \in$ null N then from the basis given in theorem 10.7.1 (8D), we can represent v as $v = \sum_{0 \le i \le m_j, 1 \le j \le n} \alpha_{i,j} N^i v_j$ so $0 = Nv = \sum_{1 \le i \le m_j} \alpha_{i,j} N^i v_j$. We follows $\alpha_{i,j} = 0$ for all $1 \le j \le n, 1 \le i \le m_j$, which means $v = \sum_{j=1}^{n} \alpha_{m_j,j} N^{m_j} v_j$. Thus, $N^{m_1}v_1, \dots, N^{m_n}v_n$ is basis of null N.

exer:8D:7

7. From theorem 10.7.3 (8D), draw out the Jordan matrix for T that makes p the minimal polynomial and q the characteristic polynomial. After that, we can choose the standard basis $e_1, \dots, e_n$ as the Jordan basis for this matrix. This will define an operator T that satisfies the condition.

exer:8D:8 8. If there does not exist a direct sum decomposition of V into two proper subspaces invariant under T then from theorem 10.3.1 (8B), T must have only one eigenvalue λ and $V = G(\lambda, T) = \text{null } (T - \lambda I)^{\dim V}$. We follow $(T - \lambda I)$ is a nilpotent operator on V. If $\dim \text{null } (T-\lambda I) \geq 2$ then according to theorem 10.7.1 (8D), for $n = \dim \text{null } (T-\lambda I) \geq 2$, V has a Jordan basis $(T - \lambda I)^{m_1}v_1, \ldots, v_1, \ldots, (T - \lambda I)^{m_n}v_n, \ldots, v_n$ for T. Let $U = \text{span } ((T-\lambda I)^{m_1}v_1, \ldots, v_1)$ and $W = \text{span } ((T-\lambda I)^{m_2}v_2, \ldots, v_2, \ldots, (T-\lambda I)^{m_n}v_n, \ldots, v_n)$ then $V = U \oplus W$. If $u \in U$ then $Tu = (T-\lambda I)u + \lambda u \in U$ so U is invariant under T. Similarly, W is invariant under T. This contradicts to the original assumption. Hence, we must have $\dim \text{null } (T-\lambda I) = 1$, which again from theorem 10.7.1 (8D), $(T-\lambda I)^{\dim V - 1}v, \ldots, v$ is a basis of V. This follows the smallest d so $(T - \lambda I)^d = 0$ is $d = \dim V$. We find $(z - \lambda)^{\dim V}$ is the minimal polynomial of T.

   If $(z-\lambda)^{\dim V}$ is the minimal polynomial of T then there exists v so $(T-\lambda I)^{n-1}v, \ldots, v$ is a basis of V with $n = \dim V$. Assume the contrary that V has a direct sum decomposition into two proper subspaces U, W invariant under T. If $u \in U$, write $u = \sum_{i=0}^{n-1} a_i(T-\lambda I)^i v$ where $a_i \in \mathbf{C}$ for $0 \leq i \leq n-1$. Let a_i be the first nonzero number in $a_0, \ldots, a_{n-1}$. Since U is invariant under T so U is also invariant under $T - \lambda I$, which means $(T-\lambda I)^{n-i-1}u = a_i(T - \lambda I)^{n-1}v \in U$. Therefore, $(T - \lambda I)^{n-1}v \in U$. We can apply similar method to W can conclude that $(T-\lambda I)^{n-1}v \in W$. Hence, $U \cap W \neq \{0\}$, so we can't have $V = U \oplus W$. Thus, there does not exist a direct sum decomposition of V into two proper subspaces invariant under T.

11. Chepter 9: Operators on Real Vector Spaces

11.1. 9A: Complexification

Definition 11.1.1 (Complexification $V_{\mathbf{C}}$ of V)**.** The complexification $V_{\mathbf{C}}$ of V is $V \times V$, where each element is of the form (u, v) or $u + iv$ for $u, v \in V$. We define addition on $V_{\mathbf{C}}$ as $(u_1 + iv_1) + (u_2 + iv_2) = (u_1 + v_1) + i(u_2 + v_2)$ for $u_i, v_i \in V$ and scalar multiplication as $(a + ib)(u + iv) = (au - bv) + i(by + av)$ for $a, b \in \mathbf{R}$ and $u, v \in V$.

Construction of $V_{\mathbf{C}}$ from V can be thought as generalizing the construction of $\mathbf{C}^n$ from $\mathbf{R}^n$.

Theorem 11.1.2

Complexification of V is a complex vector space. Any basis of V is a basis of $V_{\mathbf{C}}$. Even better, any subspace of $V_{\mathbf{C}}$ has a basis containing vectors in V.

Definition 11.1.3 (Complexification of operator T)**.** For real vector space V, complexification $T_{\mathbf{C}}$ of any $T \in \mathcal{L}(V)$ is the operator $T_{\mathbf{C}} \in \mathcal{L}(V_{\mathbf{C}})$ such that $T_{\mathbf{C}}(u + iv) = Tu + iTv$ for any $u, v \in V$.

lexification_property **Theorem 11.1.4** An operator T on a real vector space V has minimal/characteristic polynomial same as its complexification $T_{\mathbf{C}}$. This means that:

1. Minimal/characteristic polynomials of $T_{\mathbf{C}}$ have real coefficients.
2. Real zeros of the characteristic polynomial are eigenvalues of T.
3. Multiplicity of $\lambda \in \mathbf{C}$ as eigenvalue of $T_{\mathbf{C}}$ equals the multiplicity of $\overline{\lambda}$ as eigenvalue of $T_{\mathbf{C}}$. In other words $\dim G(\lambda, T_{\mathbf{C}}) = \dim G(\overline{\lambda}, T_{\mathbf{C}})$.
4. If $\dim V$ is odd then T has at least one eigenvalue.

11.2. Exercises 9A

exer:9A:1 1. Done.

exer:9A:2 2. Show $T_{\mathbf{C}} \in \mathcal{L}(V_{\mathbf{C}})$. For $u_1 + iv_1, u_2 + iv_2 \in V_{\mathbf{C}}$ then

$$T_{\mathbf{C}}(u_1 + iv_1 + u_2 + iv_2) = T(u_1 + u_2) + iT(v_1 + v_2) = T_{\mathbf{C}}(u_1 + iv_1) + T_{\mathbf{C}}(u_2 + iv_2).$$

Similarly, $T_{\mathbf{C}}(\lambda(u + iv)) = \lambda T_{\mathbf{C}}(u + iv)$.

exer:9A:3 3. Not hard.

exer:9A:4 4. Not hard.

exer:9A:5 5. $(S+T)_{\mathbf{C}}(u+iv) = (S+T)u + i(S+T)v = (S_{\mathbf{C}} + T_{\mathbf{C}})(u+iv)$.

exer:9A:6 6. Prove injectivity and surjectivity.

exer:9A:7 7. Since $(N_{\mathbf{C}})^n(u+iv) = N^n u + iN^n v$ so N is nilpotent iff $N_{\mathbf{C}}$ is nilpotent.

exer:9A:8 8. This follows $5, 7$ are eigenvalues of $T_{\mathbf{C}}$ and since characteristic polynomial of $T_{\mathbf{C}}$ of degree 3 so $T_{\mathbf{C}}$ has no nonreal eigenvalues.

exer:9A:9 9. Since T is an operator on odd-dimensional real vector space so T has a real eigenvalue. On the other hand, since $T^2 + T + I$ is nilpotent so minimal polynomial $p(T)$ of T divides $(T^2 + T + I)^7$. Since minimal polynomial $p(T)$ of T has real coefficients so $p(T) = (T^2 + T + I)^k$ for some $k < 7$. However, this implies $p(T)$ has no real roots, which means T has no real eigenvalue according to 10.5.5, a contradiction to previous claim. Thus, there does not exists $T \in \mathcal{L}(\mathbf{R}^7)$ so $T^2 + T + I$ is a nilpotent.

exer:9A:10 10. The difference between this exercise and the previous exercise 9 is that minimal polynomial of T need not to have real coefficients. Since $T^2 + T + I$ is nilpotent, we aim to construct T so its minimal polynomial is of the form $p(T) = (T^2+T+I)^k = \left(T - \frac{-1+i\sqrt{3}}{2}I\right)^k \left(T - \frac{-1-i\sqrt{3}}{2}I\right)^k$ where $k < 7/2$ because minimal polynomial must have degree at most $\dim \mathbf{C}^7 = 7$. Say

$k=3$. From example 10.7.4, let $\lambda_+ = \frac{-1+i\sqrt{3}}{2}$ and $\lambda_- = \frac{-1-i\sqrt{3}}{2}$, we can construct a Jordan matrix as follow:

$$\mathcal{M}(T) = \begin{pmatrix} \lambda_+ & 1 & & & & & \\ & \lambda_+ & 1 & & & & \\ & & \lambda_+ & & & & \\ & & & \lambda_+ & & & \\ & & & & \lambda_- & 1 & \\ & & & & & \lambda_- & 1 \\ & & & & & & \lambda_- \end{pmatrix}$$

WLOG, say that $(0,\dots,0,1,0,\dots,0)$ is Jordan basis of T then from the above matrix we can construct T.

exer:9A:11 11. The minimal polynomial p of T must divide x^2+bx+c and since eigenvalue λ of T are roots of p so T has real eigenvalue λ if λ is root of x^2+bx+c, i.e. $b^2 \geq 4c$.

exer:9A:12 12. Similarly to exercise 11, minimal polynomial p of T must divide $(z^2+bz+c)^{\dim V}$, and since $(z^2+bz+c)^{\dim V}$ has no real root, p has no real root, i.e. T has no eigenvalues.

exer:9A:13 13. Since $W = \text{null } (T^2+bT+cI)^j$ is a vector space so $T^2+bT+cI$ is nilpotent on W, which according to previous exercise 12 implies that T has no eigenvalues. This follows $\dim W$ is even.

exer:9A:14 14. T^2+T+I is nilpotent implies $T_{\mathbf{C}}$ has exactly two eigenvalues λ and $\overline{\lambda}$ with same multiplicity according to theorem 11.1.4. Therefore, the characteristic polynomial of $T_{\mathbf{C}}$ is $p(z) = (z^2+z+1)^k$. Since $\dim V = 8$ so $\deg p = 8$ or $k=4$. Thus, the characteristic polynomial of T is $(z^2+z+1)^4$ so $(T^2+T+I)^4 = 0$.

exer:9A:15 15. Let U be an invariant subspace of V. Since T has no eigenvalue so $T|_U \in \mathcal{L}(U)$ has no eigenvalue. According to theorem 11.1.4, U has even dimension.

exer:9A:16 16. If $T^2 = -I$ then T has no eigenvalue according to exercise 12, which implies V has even dimension according to theorem 11.1.4.

Conversely, if V has even dimension. Inspired from example 7.1.1, for any basis $a_1, b_1, a_2, b_2, \dots, a_n, b_n$ of V, let $T \in \mathcal{L}(V)$ such that $T(a_i) = b_i, T(b_i) = -a_i$. With this, we can see that $T^2+I = 0$.

exer:9A:17 17. (a) We already know that V is a real vector space. For any $a,b,c,d \in \mathbf{R}$ and any $v,u \in V$ then $(a+ib)v = av+bTv \in V$ and

$$\begin{aligned} (a+bi)[(c+di)v] &= (a+bi)(cv+dTv), \\ &= (acv+bcTv)+(adTv+bdT^2v), \\ &= (ac-bd)v+(bc+ad)Tv, \\ &= (ac-bd+i(ad+bc))v, \\ &= [(a+bi)(c+di)]v. \end{aligned}$$

We also have $(a+bi)(u+v) = a(u+v) + bT(u+v) = (au+bTu) + (av+bTv) = (a+bi)u + (a+bi)v$ and similarly $(a+bi+c+di)u = (a+bi)u + (c+di)u$. Thus, V is a complex vector space.

(b) Let $\dim V = 2n$. Pick $a_i \in V$ inductively such that $a_{i+1} \notin \text{span}\,(a_1, Ta_1, a_2, Ta_1, \dots, a_i, Ta_i)$. We will first show that $a_1, Ta_1, \dots, a_n, Ta_n$ is linearly independent.

Assume the contrary that the list is linearly dependent, then there exists i such that $Ta_{i+1} \in \text{span}\,(a_1, Ta_1, \dots, a_i, Ta_i, a_{i+1})$. Let $Ta_{i+1} = \sum_{j=1}^{i}(\alpha_i a_i + \beta_i Ta_i) + ca_{i+1}$ where $\alpha_i, \beta_i \in \mathbf{R}$. Since $T^2 = -I$ so

$$-a_{i+1} = T^2 a_{i+1} = \sum_{j=1}^{i}(\alpha_i Ta_i - \beta_i a_i) + cTa_{i+1}.$$

Substituting Ta_{i+1} into the above to obtain $a_{i+1} \in \text{span}\,(a_1, Ta_1, \dots, a_i, Ta_i)$, a contradiction to our condition. Thus, $a_1, Ta_2, \dots, a_n, Ta_n$ must be linearly independent and therefore it is a basis of real vector space V.

We prove that with such definition of complex scalar multiplication then $a_1, \dots, a_n$ is the basis of V as complex vector space. Since $a_j \in V$ so $ia_j = Ta_j \in V$. Therefore, $a_1, Ta_1, \dots, a_n, Ta_n \in V$ which implies that $a_1, \dots, a_n$ spans V. The list is also linear independent because $\sum_{i=1}^{n}(\alpha_i + i\beta_i)a_i = \sum_{i=1}^{n} \alpha_i a_i + \beta_i Ta_i$.

Thus, the dimension of V as a complex vector space is half the dimension of V as real vector space.

exer:9A:18 18. $(b) \implies (a)$: If there exists a basis of V wrt to which T has an upper-triangular matrix, then it is also a basis of $V_{\mathbf{C}}$ with respect to which $T_{\mathbf{C}}$ has an upper-triangular matrix. Therefore, from theorem 7.3.4 (5B), we find $T_{\mathbf{C}}$ has real eigenvalues.

$(a) \implies (c)$: If all eigenvalues of $T_{\mathbf{C}}$ are real. Consider $G(\lambda, T_{\mathbf{C}})$ where $\lambda \in \mathbf{R}$ is an eigenvalue of $T_{\mathbf{C}}$. Consider basis $x_1 + iy_1, \dots, x_n + iy_n$ of $G(\lambda, T_{\mathbf{C}})$ where $x_j, y_j \in V$. We have $T_{\mathbf{C}}(x_j + iy_j) = Tx_j + iTy_j = \lambda(x_j + iy_j)$ which implies $Tx_j = \lambda x_j, Ty_j = \lambda y_j$ for all $1 \le j \le n$. On the other hand, note that $x_j + iy_j \notin U = \text{span}\,(x_1 + iy_1, \dots, x_{j-1} + iy_{j-1})$ so at least one of x_j, y_j must not be in U. From this, we can inductively choose either x_i or y_i to create a basis of $G(\lambda, T_{\mathbf{C}})$ that are all in V.

From theorem 10.3.1 (8B), since $V_{\mathbf{C}} = \bigoplus_{i=1}^{m} G(\lambda_i, T_{\mathbf{C}})$, we can construct a basis of $V_{\mathbf{C}}$ consisting generalized eigenvectors that are in V. This follows V has a basis consisting of generalized eigenvectors of T.

$(c) \implies (a)$: If there is a basis of V consisting of generalized eigenvectors of T then $V = \bigoplus_{i=1}^{m} G(\lambda_i, T)$ where $\lambda_i \in \mathbf{R}$ are eigenvalues of T.

If $v_1, \dots, v_n$ is basis of $G(\lambda, T)$ then since $\lambda \in \mathbf{R}$, we have $(T_{\mathbf{C}} - \lambda I)^k(u + iv) = 0$ when $u, v \in G(\lambda, T)$. Therefore, $v_1, \dots, v_n$ is a basis of $G(\lambda, T_{\mathbf{C}})$. Hence, we obtain $V_{\mathbf{C}} = \bigoplus_{i=1}^{m} G(\lambda_i, T_{\mathbf{C}})$. This follows that $T_{\mathbf{C}}$ does not have any nonreal eigenvalue.

$(a) \implies (c) \implies (b)$: If all eigenvalues of $T_{\mathbf{C}}$ are real and then there exists a basis of V consisting of generalized eigenvectors of T. Hence, $V = \bigoplus_{i=1}^{m} G(\lambda_i, T)$ where $\lambda_i \in \mathbf{R}$ are eigenvalues of T. With this and from the proof of theorem 10.7.2 (8D), we find that there is a Jordan basis for T of V. This proves (b).

exer:9A:19 19. Bring back to working with complex vector space. Since null $T^{n-2} \neq$ null T^{n-1} so null $T_{\mathbf{C}}^{n-2} \neq$ null $T_{\mathbf{C}}^{n-1}$. Hence, from exercise 4, we find $T_{\mathbf{C}}$ has at most 2 eigenvalues, one of which must be 0. This follows the other must be real eigenvalue. Therefore, T has at most 2 eigenvalues and $T_{\mathbf{C}}$ has no nonreal eigenvalues.

11.3. 9B: Operators on Real Inner Product Spaces

theo:9.34:9B **Theorem 11.3.1** (Characterization of normal vectors when $\mathbf{F} = \mathbf{R}$) Suppose V is a real inner product space and $T \in \mathcal{L}(V)$. Then the following are equivalent:

1. T is normal.
2. There is an orthonormal basis of V with respect to which T has a block diagonal matrix such that each block is a 1-by-1 matrix or a 2-by-2 matrix of the form $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ with $b > 0$.

By looking at the characteristic polynomial of normal operator T, one can count number of 2-by-2 block diagonal matrices.

Using this theorem 11.3.1 and theorem 9.1.4, one can easily prove the Real Spectral Theorem 9.3.2.

theo:9.36:9B **Theorem 11.3.2** (Description of isometries when $\mathbf{F} = \mathbf{R}$) Suppose V real inner product space and $S \in \mathcal{L}(V)$. Then the following are equivalent:

1. S is an isometry.
2. There is an orthonormal basis of V with respect to which S has a block diagonal matrix such that each block on the diagonal is a 1-by-1 matrix containing 1 or -1 or is a 2-by-2 matrix of the form $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ with $\theta \in (0, 2\pi)$.

We can also define an inner product for complexification $V_{\mathbf{C}}$ as indicated in exercises 3, i.e.

$$\langle u + iv, x + iy \rangle = \langle u, x \rangle + \langle v, y \rangle + (\langle v, x \rangle - \langle u, y \rangle)\, i.$$

11.4. Exercises 9B

exer:9B:1 1. From theorem 11.3.2, S has block diagonal 3-by-3 matrix so at least one block on the diagonal is a 1-by-1 matrix. If it's the first block then pick $x = (1, 0, 0)$, if it's the last block then $x = (0, 0, 1)$.

exer:9B:2 2. Obviously true from theorem 11.3.2.

exer:9B:3 3. Jump back to the definition of inner product.

Check positivity and definiteness: $\langle u+iv, u+iv\rangle = \langle u,u\rangle + \langle v,v\rangle \geq 0$ and is 0 when $u=v=0$.

Check additivity in first slot and homogeneity in first slot.

Check conjugate symmetry:

$$\begin{aligned} \overline{\langle u+iv, x+iy\rangle} &= \langle u,x\rangle + \langle v,y\rangle - (\langle v,x\rangle - \langle u,y\rangle)\, i, \\ &= \langle x,u\rangle + \langle y,v\rangle + (\langle y,u\rangle - \langle x,v\rangle)\, i, \\ &= \langle x+iy, u+iv\rangle . \end{aligned}$$

exer:9B:4 4. It suffices to show $\langle T_{\mathbf{C}}(u+iv), x+iy\rangle = \langle u+iv, T_{\mathbf{C}}(x+iy)\rangle$.

exer:9B:5 5. If T is self-adjoint then $T_{\mathbf{C}}$ is self-adjoint according to exercise 4, which means all eigenvalues of $T_{\mathbf{C}}$ are real according to theorem 9.1.5. By Complex Spectral Theorem 9.3.1, $V_{\mathbf{C}}$ has orthonormal basis consisting of eigenvectors of $T_{\mathbf{C}}$. This implies $V_{\mathbf{C}} = \bigoplus_{i=1}^m E(\lambda_i, T_{\mathbf{C}})$. In each subspace $E(\lambda_i, T_{\mathbf{C}})$, there exists a basis containing vectors in V. By applying Gram-Schmidt procedure 8.3.2 to this basis, we obtain an orthonormal basis of $E(\lambda, T_{\mathbf{C}})$ containing vectors in V. Combining all these vectors and the fact that eigenvectors corresponding to distinct eigenvalues of normal operator are orthorgonal 9.1.7, we obtain an orthonormal basis for V containing all eigenvectors of T.

exer:9B:6 6. Look at the proof and write out the matrix so B is nonzero matrix. Say $\mathcal{M}(T) = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ wrt to standard basis where $T \in \mathcal{L}(\mathbf{R}^2)$. Then $U = \text{span}\ (1,0)$ is invariant under T but not $U^{\perp} = \text{span}\ (0,1)$.

exer:9B:7 7. Let such basis of V be $v_1, \ldots, v_n$ and show that $Tv_i = (T_1 \cdots T_m)v_i$.

exer:9B:8 8. According to exercise 21 (7A), the list

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \ldots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \cdots, \frac{\sin nx}{\sqrt{\pi}}$$

is the orthonormal basis of V, and the matrix of D with this respect to this basis will give the desired form in 11.3.1. In particular, $U_i = \text{span}\ \left(\frac{\cos ix}{\sqrt{\pi}}, \frac{\sin ix}{\sqrt{\pi}}\right)$ is invariant under T where $\mathcal{M}(T|_{U_i}) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

12. Chapter 10: Trace and Determinant

12.1. 10A: Trace

prop:10.4:10A **Proposition 12.1.1** (Matrix of product of linear maps) Suppose $u_1, \dots, u_n$ and $v_1, \dots, v_n$ and $w_1, \dots, w_n$ are bases of V. Suppose $S, T \in \mathcal{L}(V)$. Then

$$\begin{aligned}\mathcal{M}(ST, (u_1, \dots, u_n), (w_1, \dots, w_n)) =& \mathcal{M}(S, (v_1, \dots, v_n), (w_1, \dots, w_n)) \\ & \mathcal{M}(T, (u_1, \dots, u_n), (v_1, \dots, v_n)).\end{aligned}$$

coroll:10.7:10A **Corollary 12.1.2** (Change of basis) Suppose $T \in \mathcal{L}(V)$. Let $u_1, \dots, u_n$ and $v_1, \dots, v_n$ be bases of V.
Let $A = \mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n))$ then

$$\mathcal{M}(T, (u_1, \dots, u_n)) = A^{-1} \mathcal{M}(T, (v_1, \dots, v_n)) A.$$

theo:10.16:10A **Theorem 12.1.3** (Trace of an operator equals to trace of its matrix) For $T \in \mathcal{L}(V)$ then trace T = trace $\mathcal{M}(T)$ = (trace $\mathcal{M}(T_{\mathbf{C}})$). In other words, for any basis of T, the sum of diagonal entries of $\mathcal{M}(T)$ is constant and is equal to sum of eigenvalues of T (of $T_{\mathbf{C}}$ if V is real vector space), with each eigenvalue repeated according to its multiplicity.

12.2. Exercises 10A

exer:10A:1 1. If $A = \mathcal{M}(T, v_1, \dots, v_n))$ is invertible there exists inverse B of A. We define an operator $S \in \mathcal{L}(V)$ such that $\mathcal{M}(S, (v_1, \dots, v_n)) = B$. By using proposition 12.1.1, we can show that B is inverse of T.

   If T is invertible, then there exists inverse T^{-1} of T. Hence, using proposition 12.1.1 to matrices $\mathcal{M}(TT^{-1}, (v_1, \dots, v_n))$ and $\mathcal{M}(T^{-1}T, (v_1, \dots, v_n))$ we obtain that $\mathcal{M}(T, (v_1, \dots, v_n))$ is invertible.

exer:10A:2 2. Bring back to woring with operators. Define $S, T \in \mathcal{L}(V)$ so $\mathcal{M}(S, (v_1, \dots, v_n)) = A, \mathcal{M}(T, (v_1, \dots, v_n)) = B$ then $ST = I$ as $\mathcal{M}(ST) = I$. Applying exercise 10, we implies $TS = I$ so $BA = I$.

exer:10A:3 3. Consider an arbitrary basis $v_1, \dots, v_n$ of V. Let $Tv_1 = \sum_{i=1}^n \alpha_i v_i$. Since $\mathcal{M}(T, (v_1, \dots, v_n)) = \mathcal{M}(T, (2v_1, \dots, v_n))$ so $\alpha_i = 2\alpha_i$ for all $i \neq 1$ which implies $\alpha_i = 0$ for all $i \neq 1$. Hence, $Tv_1 = \alpha_1 v_1$. Similarly, we obtain $Tv_j = \alpha_j v_j$ for all j.

   On the other hand, since $\mathcal{M}(T, (v_1, v_2, \dots, v_n)) = \mathcal{M}(T, (v_2, v_1, \dots, v_n))$ so $\alpha_1 = \alpha_2$. As a result, we obtain $T = \alpha I$.

exer:10A:4 4. Since $Tv_k = u_k$ so $\mathcal{M}(T, (v_1, \dots, v_n), (u_1, \dots, u_n)) = I$. Therefore, by applying proposition 12.1.1, we obtain

$$\mathcal{M}(T, (v_1, \dots, v_n)) = \mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n)).$$

exer:10A:5 5. Bring back to working with operator. Let $T \in \mathcal{L}(V)$ be an operator so $\mathcal{M}(T,(v_1,\ldots,v_n)) = B$. Since V is complex vector space so according to theorem 10.3.3 (or 10.7.2), there exists basis $u_1,\ldots,u_n$ so $\mathcal{M}(T,(u_1,\ldots,u_n))$ is an upper-triangular matrix. Now applying corollary 12.1.2 and we are done.

exer:10A:6 6. Let $T \in \mathcal{L}(\mathbf{R}^2)$ so $\mathcal{M}(T) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Then trace $(T^2) = -2 < 0$.

exer:10A:7 7. trace $(T^2) = \sum_{i=1}^m \lambda_i^2 \geq 0$.

exer:10A:8 8. If $w = 0$ then trace $T = 0$. If $w \neq 0$ then choose basis of V contatining w, with this, we obtain trace $T = \langle w, v\rangle$.

exer:10A:9 9. If $P^2 = P$ then according to exercise 7 (8C), the characteristic polynomial of P (or $P_{\mathbf{C}}$) is $z^m(z-1)^n$ where $n = \dim \operatorname{range} P = \dim \operatorname{range} P_{\mathbf{C}}$. This follows trace $T = \dim \operatorname{range} P$ according to the definition of trace of an operator.

exer:10A:10 10. For orthonormal basis $v_1,\ldots,v_n$ of V, then according to theorem 9.1.4 (7A), $\mathcal{M}(T,(v_1,\ldots,v_n))$ is conjugate transpose of $\mathcal{M}(T^*,(v_1,\ldots,v_n))$. Therefore, trace $T^* = \overline{\text{trace } T}$ according to theorem 12.1.3 (10A)

exer:10A:11 11. From theorem 9.5.2, if T is positive then all eigenvalues of T are nonnegative. As trace T is sum of all eigenvalues and trace $T = 0$, we obtain all eigenvalues are 0. Thus, $T = 0$.

exer:10A:12 12. We have trace (PQ) = trace $\mathcal{M}(PQ)_{\mathbf{C}}$ = trace $(\mathcal{M}(P_{\mathbf{C}})\mathcal{M}(Q_{\mathbf{C}}))$. With this, we can choose basis for $P_{\mathbf{C}}$ and $Q_{\mathbf{C}}$ as in theorem 10.3.3 (8B) so $\mathcal{M}(P_{\mathbf{C}})$ and $\mathcal{M}(Q_{\mathbf{C}})$ are upper-triangular matrices. Since P,Q are orthogonal projections then so are $P_{\mathbf{C}}, Q_{\mathbf{C}}$. Hence, $P_{\mathbf{C}}^2 = P_{\mathbf{C}}$ and $Q_{\mathbf{C}}^2 = Q_{\mathbf{C}}$ according to theorem 8.5.5 (6C). This implies that eigenvalues of $P_{\mathbf{C}}$ and $Q_{\mathbf{C}}$ are only 0 and 1. This means trace $\mathcal{M}(P_{\mathbf{C}})\mathcal{M}(Q_{\mathbf{C}}) \geq 0$, as desired.

exer:10A:13 13. $51 - 40 + 1 = 12$ and $12 - 24 - (-48) = 36$ so the third eigenvalue if 36.

exer:10A:14 14. We have trace (cT) = trace $\mathcal{M}(cT)$ = trace $c\mathcal{M}(T)$ = $c$trace T.

exer:10A:15 15. We have

$$\begin{aligned}
\text{trace } (ST) &= \text{trace } \mathcal{M}(ST), \\
&= \text{trace } (\mathcal{M}(S)\mathcal{M}(T)), \\
&= \text{trace } (\mathcal{M}(T)\mathcal{M}(S)), \\
&= \text{trace } \mathcal{M}(TS), \\
&= \text{trace } TS.
\end{aligned}$$

exer:10A:16 16. Let $ST = T = S = I$ with $\dim V = 2$ then trace S = trace T = trace $ST = 2$ and therefore, trace $ST \neq$ trace Strace T.

exer:10A:17 17. Work with $\mathcal{M}(T)$ and $\mathcal{M}(S)$. Choose $S_{i,j} \in \mathcal{L}(V)$ so $\mathcal{M}(S_{i,j})_{i,j} = 1$ while the rest entries are 0. Hence, this implies $\mathcal{M}(T)_{j,i} = 0$. Thus, this concludes $T = 0$.

exer:10A:18 18. For orthonormal basis $e_1, \dots, e_m$ of V then from theorem 9.1.4 (7A), $\mathcal{M}(T, (e_1, \dots, e_m))$ is the conjugate transpose of $\mathcal{M}(T^*, (e_1, \dots, e_m))$. This means, the i-th row of $\mathcal{M}(T^*)$ is the conjugate of the i-th column of $\mathcal{M}(T)$. Hence, $\|Te_i\|^2$ equals the inner product of these two vectors, which is the $(i, i$-th entry of $\mathcal{M}(T^*T, (e_1, \dots, e_m))$. Thus, trace $(T^*T) = \sum_{i=1}^m \|Te_i\|^2$. This implies that then sum is independent from the choice of orthonormal basis.

exer:10A:19 19. From previous exercise, we have $\langle T, T \rangle \geq 0$ and is 0 when when trace $TT^* = 0$, which follows $\|T^*e_i\| = 0$ for orthonormal basis $e_1, \dots, e_m$ of V. This implies $T^* = 0$ so $T = 0$. Thus, positivity and definiteness are proven.

   Addivitity in first slot is true because of the distributive property for matrix addition and matrix multiplication 13 (3C).

   Homogeneity in first slot is true using exercise 14 (10A).

   Conjugate Symmetry is true because ST^* is adjoint of TS^*, which according to exercise 10 (10A) then

$$\langle S, T \rangle = \text{trace } (ST^*) = \overline{\text{trace } (TS^*)} = \overline{\langle T, S \rangle}.$$

exer:10A:20 20. Let $e_1, \dots, e_n$ be orthonormal basis of V which gives the desired matrix. From exercise 18 (10A), we find

$$\text{trace } (T^*T) = \sum_{i=1}^n \|Te_i\|^2 = \sum_{k=1}^n \sum_{j=1}^n |A_{j,k}|^2.$$

   It suffices to show $\sum_{i=1}^n |\lambda_i|^2 \leq \text{trace } (T^*T)$. We go and pick different basis of V. Pick orthonormal basis $e_1, \dots, e_n$ as in theorem 10.3.3 (8B) so $\mathcal{M}(T)$ is an upper triangular matrix where all the eigenvalues are on its diagonal. Hence, $\|Te_i\|^2 \geq |\lambda_i|^2$ for all $1 \leq i \leq n$. Combining with exercise 18 (10A), we obtain the desired inequality.

exer:10A:21 21. Applying exercise 18 (10A), for any orthonormal basis $e_1, \dots, e_n$ of V so $e_1 = v$, we have

$$\text{trace } (TT^*) = \sum_{i=1}^n \|T^*e_i\|^2 \leq \sum_{i=1}^n \|Te_i\|^2 = \text{trace } (T^*T).$$

   On the other hand, according to exercise 15 (10A), we find trace $(TT^*) = \text{trace } (T^*T)$. Thus, the equality holds when $\|T^*e_i\| = \|Te_i\|$ so $\|T^*v\| = \|Tv\|$. Since this is true for any $v \in V$, we obtain that T is normal

13. Summary

Example 13.0.1 (Inner product space of continuous real-valued functions)

Let $\mathbf{C}[-\pi,\pi]$ be the vector space of continuous real-valued functions on $[-\pi,\pi]$ with inner product

$$\langle f, g\rangle = \int_{-\pi}^{\pi} f(x)g(x)dx.$$

For some positive integer integer n then

1. (Exercise 4 (6B)) $\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \dots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots, \frac{\sin nx}{\sqrt{\pi}}$ is an orthonormal list of vectors in $\mathbf{C}[-\pi,\pi]$.

2. Let $V = \operatorname{span}(1, \cos x, \cos 2x, \dots, \cos nx, \sin x, \sin 2x, \dots, \sin nx)$ be a subspace of $\mathbf{C}[-\pi,\pi]$.

   a) Define $D \in \mathcal{L}(V)$ by $Df = f'$. Then $D^* = -D$ which implies D is normal but not self-adjoint (exercise 21 (7A)). From previous remark, we can also find an orthonormal basis of V such that the matrix of D has the form described in theorem 11.3.1 (see exercise 8 (9B)).

   b) Define $T \in \mathcal{L}(V)$ by $Tf = f''$. Then T is self-adjoint and $-T$ is positive operator (exercise 14 (7C)).

Example 13.0.2 (Projection operator)

A linear operator P on vector space V is a projection on V if $P^2 = P$. That is, it leaves its image unchanged.

1. $V = \operatorname{null} P \oplus \operatorname{range} P$ (exercise 4(5B)).

2. Characteristic polynomial of P is $z^m(z-1)^n$ where $m = \dim \operatorname{null} P$ and $n = \dim \operatorname{range} P$ (exercise 7 (8C)).

3. $\operatorname{trace} P = \dim \operatorname{range} P$ (exercise 9 (10A)).

4. P is an orthogonal projection 8.5.3 iff it is self-adjoint (exercise 11 (7A)).

All norms are equivalent?? https://math.stackexchange.com/questions/112985/every-linear-mapping 833995#833995

All we just learned is about **normed vector space**, i.e. vector space over the real or complex numbers, on which a norm is defined. See wiki for more.

Two norms are equivalent on finite dimensional space?? See exercise 29 (6A).

Every inner product space is a normed space, but not every normed space is an inner product space. See here.

The motivation for the equivalence of norms on V come from the fact that we want the two norms to define the same open subsets of V. See here for more. See here for proofs about norms being equivalent in finite dimensional normed space.

14. Interesting problems

1. How many there are k-dimensional subspaces in n-dimensional vector space V over $\mathbf{F}_p$ for p is a prime. (See here)

   Proof. So there are $(p^n-1)(p^n-p)\cdots(p^n-p^{k-1})$ linearly independent lists of length k in V and since each k-dimensional subspace of V contributes $(p^k-1)(p^k-p)\cdots(p^k-p^{k-1})$ such lists, so there are $\frac{(p^n-1)(p^n-p)\cdots(p^n-p^{k-1})}{(p^k-1)(p^k-p)\cdots(p^k-p^{k-1})}$ k-dimensional subspaces of V. □

2. (AoPS) Let $f \in \mathbb{Q}[X]$ be a polynomial of degree $n > 0$.Let $p_1, p_2, .., p_{n+1}$ be distinct prime numbers.Show that there exists a non-zero polynomial $g \in \mathbb{Q}[X]$ such that $fg = \sum_{i=1}^{n+1} c_i X^{p_i}$ with $c_i \in \mathbb{Q}$.

   Proof. Let $\mathcal{P}_n(\mathbf{Q})$ be the $\mathbf{Q}$-vector space of polynomials with degree at most n. For each i, let $X^{p_i} = fq_i + r_i$ where r_i is the remainder of X^{p_i} modulo f. Since $\deg f = n$ so $r_i \in \mathcal{P}_n(\mathbf{Q})$ for all i. On the other hand, since $p_1, \dots, p_{n+1}$ are primes, this follows $r_1, \dots, r_{n+1}$ are linearly independent in $\mathcal{P}_n(\mathbf{Q})$, which means they form a basis of $\mathcal{P}_n(\mathbf{Q})$. Therefore, as $f \in \mathcal{P}_n(\mathbf{Q})$, there exists $c_i \in \mathbf{Q}$ such that $f = \sum_{i=1}^{n+1} c_i r_i$.

   Therefore, we have

   $$\sum_{i=1}^{n+1} c_i X^{p_i} = \sum_{i=1}^{n+1} c_i(fq_i + r_i) = f\left(\sum_{i=1}^{n+1} c_i q_i + 1\right) = fg.$$

   Since $p_{n+1} > n$ so $\deg q_{n+1} \geq 1$ so $\deg g \geq 1$. □

3. (MSE) Show that any $n \times n$ marix A can be written as sum of two non-singular matrices.

   Proof. There are two ways presented in the link. □

15. New knowledge

1. Rational canonical form, AoPS

2. Difference between $\mathbf{R}$ and $\mathbf{C}$? algebraic closed field?