Solutions to Axler, Linear Algebra Done Right 2nd Ed.

Edvard Fagerholm
edvard.fagerholm@{helsinki.fi|gmail.com}

Beware of errors. I read the book and solved the exercises during spring break (one week), so the problems were solved in a hurry. However, if you do find better or interesting solutions to the problems, I'd still like to hear about them. Also please don't put this on the Internet to encourage copying homework solutions...

1 Vector Spaces

1. Assuming that $\mathbb{C}$ is a field, write $z = a + bi$. Then we have $1/z = \overline{z}/\overline{z}z = \overline{z}/|z|^2$. Plugging in the numbers we get $1/(a + bi) = a/(a^2 + b^2) - bi/(a^2 + b^2) = c + di$. A straightforward calculation of $(c+di)(a+bi) = 1$ shows that this is indeed an inverse.

2. Just calculate $((1 + \sqrt{3})/2)^3$.

3. We have $v + (-v) = 0$, so by the uniqueness of the additive inverse (prop. 1.3) $-(-v) = v$.

4. Choose $a \neq 0$ and $v \neq 0$. Then assuming $av = 0$ we get $v = a^{-1}av = a^{-1}0 = 0$. Contradiction.

5. Denote the set in question by A in each part.

   (a) Let $v, w \in A$, $v = (x_1, x_2, x_3)$, $w = (y_1, y_2, y_3)$. Then $x_1 + 2x_2 + 3x_3 = 0$ and $y_1 + 2y_2 + 3y_3 = 0$, so that $0 = x_1 + 2x_2 + 3x_3 + y_1 + 2y_2 + 3y_3 = (x_1 + y_1) + 2(x_2 + y_2) + 3(x_3 + y_3)$, so $v + w \in A$. Similarly $0 = a0 = ax_1 + 2ax_2 + 3ay_3$, so $av \in A$. Thus A is a subspace.

   (b) This is not a subspace as $0 \notin A$.

   (c) We have that $(1, 1, 0) \in A$ and $(0, 0, 1) \in A$, but $(1, 1, 0)+(0, 0, 1) = (1, 1, 1) \notin A$, so A is not a subspace.

   (d) Let $(x_1, x_2, x_3), (y_1, y_2, y_3) \in A$. If $x_1 = 5x_3$ and $y_1 = 5y_3$, then $ax_1 = 5ax_3$, so $a(x_1, x_2, x_3) \in A$. Similarly $x_1 + y_1 = 5(x_3 + y_3)$, so that $(x_1, x_2, x_3) + (y_1, y_2, y_3) \in A$. Thus A is a subspace.

6. Set $U = \mathbb{Z}^2$.

7. The set $\{(x, x) \in \mathbb{R}^2 \mid x \in \mathbb{R}\} \cup \{(x, -x) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$ is closed under multiplication but is trivially not a subspace ($(x, x) + (x, -x) = (2x, 0)$ doesn't belong to it unless $x = 0$).

8. Let $\{V_i\}$ be a collection of subspaces of V. Set $U = \cap_i V_i$. Then if $u, v \in U$. We have that $u, v \in V_i$ for all i because V_i is a subspace. Thus $au \in V_i$ for all i, so that $av \in U$. Similarly $u + v \in V_i$ for all i, so $u + v \in U$.

9. Let $U, W \subset V$ be subspaces. Clearly if $U \subset W$ or $W \subset U$, then $U \cup W$ is clearly a subspace. Assume then that $U \not\subset W$ and $W \not\subset U$. Then we can choose $u \in U \setminus W$ and $w \in W \setminus U$. Assuming that $U \cup W$ is a subspace we have $u + w \in U \cup W$. Assuming that $u + w \subset U$ we get $w = u + w - u \subset U$. Contradiction. Similarly for $u + w \in W$. Thus $U \cup W$ is not a subspace.

10. Clearly $U = U + U$ as U is closed under addition.

11. Yes and yes. Follows directly from commutativity and associativity of vector addition.

12. The zero subspace, $\{0\}$, is clearly an additive identity. Assuming that we have inverses, then the whole space V should have an inverse U such that $U + V = \{0\}$. Since $V + U = V$ this is clearly impossible unless V is the trivial vector space.

13. Let $W = \mathbb{R}^2$. Then for any two subspaces U_1, U_2 of W we have $U_1 + W = U_2 + W$, so the statement is clearly false in general.

14. Let $W = \{p \in \mathcal{P}(\mathbb{F}) \mid p = \sum_{i=0}^{n} a_i x^i,\ a_2 = a_5 = 0\}$.

15. Let $V = \mathbb{R}^2$. Let $W = \{(x, 0) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$. Set $U_1 = \{(x, x) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$, $U_2 = \{(x, -x) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$. Then it's easy to see that $U_1 + W = U_2 + W = \mathbb{R}^2$, but $U_1 \neq U_2$, so the statement is false.

2 Finite-Dimensional Vector Spaces

1. Let $u_n = v_n$ and $u_i = v_i + v_{i+1}$, $i = 1, \ldots, n-1$. Now we see that $v_i = \sum_{j=i}^{n} u_i$. Thus $v_i \in \operatorname{span}(u_1, \ldots, u_n)$, so $V = \operatorname{span}(v_1, \ldots, v_n) \subset \operatorname{span}(u_1, \ldots, u_n)$.

2. From the previous exercise we know that the span is V. As $(v_1, \ldots, v_n)$ is a linearly independent spanning list of vectors we know that $\dim V = n$. The claim now follows from proposition 2.16.

3. If $(v_1 + w, \ldots, v_n + w)$ is linearly dependent, then we can write

$$0 = a_1(v_1 + w) + \ldots + a_n(v_n + w) = \sum_{i=1}^{n} a_i v_i + w \sum_{i=1}^{n} a_i,$$

where $a_i \neq 0$ for some i. Now $\sum_{i=1}^{n} a_i \neq 0$ because otherwise we would get

$$0 = \sum_{i=1}^{n} a_i v_i + w \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} a_i v_i$$

and by the linear independence of (v_i) we would get $a_i = 0$, $\forall i$ contradicting our assumption. Thus

$$w = -\left(\sum_{i=1}^{n} a_i\right)^{-1} \sum_{i=1}^{n} a_i v_i,$$

so $w \in \operatorname{span}(v_1, \ldots, v_n)$.

4. Yes, multiplying with a nonzero constant doesn't change the degree of a polynomial and adding two polynomials either keeps the degree constant or makes it the zero polynomial.

5. $(1, 0, \ldots), (0, 1, 0, \ldots), (0, 0, 1, 0, \ldots), \ldots$ is trivially linearly independent. Thus $\mathbb{F}^\infty$ isn't finite-dimensional.

6. Clearly $\mathcal{P}(\mathbb{F})$ is a subspace which is infinite-dimensional.

7. Choose $v_1 \in V$. Assuming that V is infinite-dimensional $\operatorname{span}(v_1) \neq V$. Thus we can choose $v_2 \in V \setminus \operatorname{span}(v_1)$. Now continue inductively.

8. $(3, 1, 0, 0, 0), (0, 0, 7, 1, 0), (0, 0, 0, 0, 1)$ is clearly linearly independent. Let $(x_1, x_2, x_3, x_4, x_5) \in U$. Then $(x_1, x_2, x_3, x_4, x_5) = (3x_2, x_2, 7x_4, x_4, x_5) = x_2(3, 1, 0, 0, 0) + x_4(0, 0, 7, 1, 0) + x_5(0, 0, 0, 0, 1)$, so the vectors span U. Hence, they form a basis.

9. Choose the polynomials $(1, x, x^2 - x^3, x^3)$. They clearly span $\mathcal{P}_3(\mathbb{F})$ and form a basis by proposition 2.16.

10. Choose a basis $(v_1, \ldots, v_n)$. Let $U_i = \operatorname{span}(v_i)$. Now clearly $V = U_1 \oplus \ldots \oplus U_n$ by proposition 2.19.

11. Let $\dim U = n = \dim V$. Choose a basis $(u_1, \ldots, u_n)$ for U and extend it to a basis $(u_1, \ldots, u_n, v_1, \ldots, v_k)$ for V. By assumption a basis for V has length n, so $(u_1, \ldots, u_n)$ spans V.

12. Let $U = \operatorname{span}(u_0, \dots, u_m)$. As $U \neq \mathcal{P}_m(\mathbb{F})$ we have by the previous exercise that $\dim U < \dim \mathcal{P}_m(\mathbb{F}) = m + 1$. As $(p_0, \dots, p_m)$ is a spanning list of vectors having length $m + 1$ it is not linearly independent.

13. By theorem 2.18 we have $8 = \dim \mathbb{R}^8 = \dim U + \dim W - \dim(U \cap W) = 4 + 4 - \dim(U \cap W)$. Since $U + W = \mathbb{R}^8$, we have that $\dim(U \cap W) = 0$ and the claim follows

14. Assuming that $U \cap W = \{0\}$ we have by theorem 2.18 that

$$9 = \dim \mathbb{R}^9 = \dim U + \dim W - \dim(U \cap W) = 5 + 5 - 0 = 10.$$

    Contradiction.

15. Let $U_1 = \{(x, 0) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$, $U_2 = \{(0, x) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$, $U_3 = \{(x, x) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$. Then $U_1 \cap U_2 = \{0\}$, $U_1 \cap U_3 = \{0\}$ and $U_2 \cap U_3 = \{0\}$. Thus the left-hand side of the equation in the problem is 2 while the right-hand side is 3. Thus we have a counterexample.

16. Choose a basis $(u_1^i, \dots, u_{n_i}^i)$ for each U_i. Then the list of vectors $(u_1^1, \dots, u_{n_m}^m)$ has length $\dim U_1 + \dots + \dim U_m$ and clearly spans $U_1 + \dots + U_m$ proving the claim.

17. By assumption the list of vectors $(u_1^1, \dots, u_{n_m}^m)$ from the proof of the previous exercise is linearly independent. Since $V = U_1 + \dots + U_n = \operatorname{span}(u_1^1, \dots, u_{n_m}^m)$ the claim follows.

3 Linear Maps

1. Any $v \neq 0$ spans V. Thus, if $w \in V$, we have $w = av$ for some $a \in \mathbb{F}$. Now $T(v) = av$ for some $a \in \mathbb{F}$, so for $w = bv \in V$ we have $T(w) = T(bv) = bT(v) = bav = aw$. Thus T is multiplication by a scalar.

2. Define f by e.g.

$$f(x, y) = \begin{cases} x, & x = y \\ 0, & x \neq y \end{cases},$$

   then clearly f satisfies the condition, but is non-linear.

3. To define a linear map it's enough to define the image of the elements of a basis. Choose a basis $(u_1, \dots, u_m)$ for U and extend it to a basis $(u_1, \dots, u_m, v_1, \dots, v_n)$ of V. If $S \in \mathcal{L}(U, W)$. Choose some vectors $w_1, \dots, w_n \in W$. Now define $T \in \mathcal{L}(V, W)$ by $T(u_i) = S(u_i)$, $i = 1, \dots, m$ and $T(v_i) = w_i$, $i = 1, \dots, n$. These relations define the linear map and clearly $T_{|U} = S$.

4. By theorem 3.4 $\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$. If $u \notin \operatorname{null} T$, then $\dim \operatorname{range} T > 0$, but as $\dim \operatorname{range} T \leq \dim \mathbb{F} = 1$ we have $\dim \operatorname{range} T = 1$ and it follows that

$\dim \operatorname{null} T = \dim V - 1$. Now choose a basis $(u_1, \ldots, u_n)$ for $\operatorname{null} T$. By our assumption $(u_1, \ldots, u_n, u)$ is linearly independent and has length $\dim V$. Hence, it's a basis for V. This implies that

$$V = \operatorname{span}(u_1, \ldots, u_n, u) = \operatorname{span}(u_1, \ldots, u_n) \oplus \operatorname{span}(u) = \operatorname{null} T \oplus \{au \mid a \in \mathbb{F}\}.$$

5. By linearity $0 = \sum_{i=1}^n a_i T(v_i) = \sum_{i=1}^n T(a_i v_i) = T(\sum_{i=1}^n a_i v_i)$. By injectivity we have $\sum_{i=1}^n a_i v_i = 0$, so that $a_i = 0$ for all i. Hence $(T(v_1), \ldots, T(v_n))$ is linearly independent.

6. If $n = 1$ the claim is trivially true. Assume that the claim is true for $n = k$. If $S_1, \ldots, S_{k+1}$ satisfy the assumptions, then $S_1 \cdots S_k$ is injective by the induction hypothesis. Let $T = S_1 \cdots S_k$. If $u \neq 0$, then by injectivity of S_{k+1} we $S_{k+1}u \neq 0$ and by injectivity of T we have $TS_{k+1}u \neq 0$. Hence TS_{k+1} is injective and the claim follows.

7. Let $w \in W$. By surjectivity of T we can find a vector $v \in V$ such that $T(v) = w$. Writing $v = a_1 v_1 + \ldots + a_n v_n$ we get $w = T(v) = T(a_1 v_1 + \ldots + a_n v_n) = a_1 T(v_1) + \ldots + a_n T(v_n)$ proving the claim.

8. Let $(u_1, \ldots, u_n)$ be a basis for $\operatorname{null} T$ and extend it to a basis $(u_1, \ldots, u_n, v_1, \ldots, v_k)$ of V. Let $U := \operatorname{span}(v_1, \ldots, v_k)$. Then by construction $\operatorname{null} T \cap U = \{0\}$ and for an arbitrary $v \in V$ we have that $v = a_1 u_1 + \ldots + a_n u_n + b_1 v_n + \ldots + b_k v_k$, so that

$$T(v) = b_1 T(v_1) + \ldots + b_k T(v_k).$$

Hence $\operatorname{range} T = T(U)$.

9. It's easy to see that $(5, 1, 0, 0), (0, 0, 7, 1)$ is a basis for $\operatorname{null} T$ (see exercise 2.8). Hence $\dim \operatorname{range} T = \dim \mathbb{F}^4 - \dim \operatorname{null} T = 4 - 2 = 2$. Thus $\operatorname{range} T = \mathbb{F}^2$, so that T is surjective.

10. It's again easy to see that $(3, 1, 0, 0, 0), (0, 0, 1, 1, 1)$ is a basis of $\operatorname{null} T$. Hence, we get $\dim \operatorname{range} T = \dim \mathbb{F}^5 - \dim \operatorname{null} T = 5 - 2 = 3$ which is impossible.

11. This follows trivially from $\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$.

12. From $\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$ it follows trivially that if we have a surjective linear map $T \in \mathcal{L}(V, W)$, then $\dim V \geq \dim W$. Assume then that $\dim V \geq \dim W$. Choose a basis $(v_1, \ldots, v_n)$ for V and a basis $(w_1, \ldots, w_m)$ for W. We can define a linear map $T \in \mathcal{L}(V, W)$ by letting $T(v_i) = w_i$, $i = 1, \ldots, m$, $T(v_i) = 0$, $i = m + 1, \ldots, n$. Clearly T is surjective.

13. We have that $\dim \operatorname{range} T \leq \dim W$. Thus we get $\dim \operatorname{null} T = \dim V - \dim \operatorname{range} T \geq \dim V - \dim W$. Choose an arbitrary subspace U of V of such that $\dim U \geq \dim V - \dim W$. Let $(u_1, \ldots, u_n)$ be a basis of U and extend it to a basis $(u_1, \ldots, u_n, v_1, \ldots, v_k)$ of V. Now we know that $k \leq \dim W$, so let $(w_1, \ldots, w_m)$ be a basis for W. Define a linear map $T \in \mathcal{L}(V, W)$ by $T(u_i) = 0$, $i = 1, \ldots, n$ and $T(v_i) = w_i$, $i = 1, \ldots, k$. Clearly $\operatorname{null} T = U$.

14. Clearly if we can find such a S, then T is injective. Assume then that T is injective. Let $(v_1, \ldots, v_n)$ be a basis of V. By exercise 5 $(T(v_1), \ldots, T(v_n))$ is linearly independent so we can extend it to a basis $(T(v_1), \ldots, T(v_n), w_1, \ldots, w_m)$ of W. Define $S \in \mathcal{L}(W, V)$ by $S(T(v_i)) = v_i$, $i = 1, \ldots, n$ and $S(w_i) = 0$, $i = 1, \ldots, m$. Clearly $ST = I_V$.

15. Clearly if we can find such a S, then T is surjective. Otherwise let $(v_1, \ldots, v_n)$ be a basis of V. By assumption $(T(w_1), \ldots, T(w_n))$ spans W. By the linear dependence lemma we can make the list of vectors $(T(w_1), \ldots, T(w_n))$ a basis by removing some vectors. Without loss of generality we can assume that the first m vectors form the basis (just permute the indices). Thus $T(w_1), \ldots, T(w_m))$ is a basis of W. Define the map $S \in \mathcal{L}(W, V)$ by $S(T(w_i)) = w_i$, $i = 1, \ldots, m$. Now clearly $TS = I_W$.

16. We have that $\dim U = \dim \operatorname{null} T + \dim \operatorname{range} T \leq \dim \operatorname{null} T + \dim V$. Substituting $\dim V = \dim \operatorname{null} S + \dim \operatorname{range} S$ and $\dim U = \dim \operatorname{null} ST + \dim \operatorname{range} ST$ we get

$$\dim \operatorname{null} ST + \dim \operatorname{range} ST \leq \dim \operatorname{null} T + \dim \operatorname{null} S + \dim \operatorname{range} S.$$

Clearly $\dim \operatorname{range} ST \leq \dim \operatorname{range} S$, so our claim follows.

17. This is nothing but pencil pushing. Just take arbitrary matrices satisfying the required dimensions and calculate each expression and the equalities easily fall out.

18. Ditto.

19. Let $V = (x_1, \ldots, _n)$. From proposition 3.14 we have that

$$\mathcal{M}(Tv) = \mathcal{M}(T)\mathcal{M}(v) = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{1,1}x_1 + \ldots + a_{1,n}x_n \\ \vdots \\ a_{m,1}x_1 + \ldots + a_{m,n}x_n \end{bmatrix}$$

which shows that $Tv = (a_{1,1}x_1 + \ldots + a_{1,n}x_n, \ldots, a_{m,1}x_1 + \ldots + a_{m,n}x_n)$.

20. Clearly $\dim \operatorname{Mat}(n, 1, \mathbb{F}) = n = \dim V$. We have that $Tv = 0$ if and only if $v = 0v_1 + \ldots, 0v_n = 0$, so $\operatorname{null} T = \{0\}$. Thus T is injective and hence invertible.

21. Let e_i denote the $n \times 1$ matrix with a 1 in the ith row and 0 everywhere else. Let T be the linear map. Define a matrix $A := [T(e_1), \ldots, T(e_n)]$. Now it's trivial to verify that
$$Ae_i = T(e_i).$$
By distributivity of matrix multiplication (exercise 17) we get for an arbitrary $v = a_1 e_1 + \ldots + a_n e_n$ that
$$\begin{aligned} Av &= A(a_1e_1 + \ldots + a_ne_n) = a_1Ae_1 + \ldots + a_nAe_n \\ &= a_1T(e_1) + \ldots + a_nT(e_n) = T(a_1e_1 + \ldots + a_ne_n), \end{aligned}$$
so the claim follows.

22. From theorem 3.21 we have ST invertible $\Leftrightarrow$ ST bijective $\Leftrightarrow$ S and T bijective $\Leftrightarrow$ S and T invertible.

23. By symmetry it's sufficient to prove this in one direction only. Thus if $TS = I$. Then $ST(Su) = S(TSu) = Su$ for all u. As S is bijective Su goes through the whole space V as u varies, so $ST = I$.

24. Clearly $TS = ST$ if T is a scalar multiple of the identity. For the other direction assume that $TS = ST$ for every linear map $S \in \mathcal{L}(V)$.

25. Let $T \in \mathcal{L}(\mathbb{F}^2)$ be the operator $T(a, b) = (a, 0)$ and $S \in \mathcal{L}(\mathbb{F}^2)$ the operator $S(a, b) = (0, b)$. Then neither one is injective and hence invertible by 3.21. However, $T + S$ is the identity operator which is trivially invertible. This can be trivially generalized to spaces arbitrary spaces of dimension ≥ 2.

26. Write
$$A := \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$
Then the system of equations in a) reduces to $Ax = 0$. Now A defines a linear map from $\mathrm{Mat}(n, 1, \mathbb{F})$ to $\mathrm{Mat}(n, 1, \mathbb{F})$. What a) states is now that the map is injective while b) states that it is surjective. By theorem 3.21 these are equivalent.

4 Polynomials

1. Let $\lambda_1, \ldots, \lambda_m$ be m distinct numbers and $k_1, \ldots, k_m > 0$ such that their sum is n. Then $\prod_{i=1}^{m} (x - \lambda_i)^{k_i}$ is clearly such a polynomial.

2. Let $p_i(x) = \prod_{j \neq i}(x - z_j)$, so that $\deg p_i = m$. Now we have that $p_i(z_j) \neq 0$ if and only if $i = j$. Let $c_i = p_i(z_i)$. Define

$$p(x) = \sum_{i=1}^{m+1} w_i c_i^{-1} p_i(x).$$

Now $\deg p = m$ and p clearly $p(z_i) = w_i$.

3. We only need to prove uniqueness as existence is theorem 4.5. Let s', r' be other such polynomials. Then we get that

$$0 = (s - s')p + (r - r') \Leftrightarrow r' - r = (s - s')p$$

We know that for any polynomials $p \neq 0 \neq q$ we have $\deg pq = \deg p + \deg q$. Assuming that $s \neq s'$ we have $\deg(r' - r) < \deg(s - s')p$ which is impossible. Thus $s = s'$ implying $r = r'$.

4. Let λ be a root of p. Thus we can write $p(x) = (x - \lambda)q(x)$. Now we get

$$p'(x) = q(x) + (x - \lambda)q'(x).$$

Thus λ is a root of p' if and only if λ is a root of q i.e. λ is a multiple root. The statement follows.

5. Let $p(x) = \sum_{i=0}^{n} a_i x^i$, where $a_i \in \mathbb{R}$. By the fundamental theorem of calculus we have a complex root z. By proposition 4.10 $\overline{z}$ is also a root of p. Then z and $\overline{z}$ are roots of equal multiplicity (divide by $(x - z)(x - \overline{z})$ which is real). Thus if we have no real roots we have an even number of roots counting multiplicity, but the number of roots counting multiplicity is $\deg p$ hence odd. Thus p has a real root.

5 Eigenvalues and Eigenvectors

1. An arbitrary element of $U_1 + \ldots + U_n$ is of the form $u_1 + \ldots + u_n$ where $u_i \in U_i$. Thus we get $T(u_1 + \ldots + u_n) = T(u_1) + \ldots + T(u_n) \in U_1 + \ldots + U_n$ by the assumption that $T(u_i) \in U_i$ for all i.

2. Let $V = \cap_i U_i$ where U_i is invariant under T for all i. Let $v \in V$, so that $v \in U_i$ for all i. Now $T(v) \in U_i$ for all i by assumption, so $T(v) \in \cap U_i = V$. Thus V is invariant under T.

3. The clame is clearly true for $U = \{0\}$ or $U = V$. Assume that $\{0\} \neq U \neq V$. Let $(u_1, \ldots, u_n)$ be a basis for U and extend it to a basis $(u_1, \ldots, u_n, v_1, \ldots, v_m)$. By our assumption $m \geq 1$. Define a linear operator by $T(u_i) = v_1$, $i = 1, \ldots, n$ and $T(v_i) = v_1$, $i = 1, \ldots, m$. Then clearly U is not invariant under T.

4. Let $u \in \operatorname{null}(T-\lambda I)$, so that $T(u)-\lambda u=0$. $ST=TS$ gives us
$$0=S(Tu-\lambda u)=STu-\lambda Su=TSu-\lambda Su=(T-\lambda I)Su,$$
so that $Su \in \operatorname{null}(T-\lambda I)$.

5. Clearly $T(1,1)=(1,1)$, so that 1 is an eigenvalue. Also $T(-1,1)=(1,-1)=-(-1,1)$ so -1 is another eigenvalue. By corollary 5.9 these are all eigenvalues.

6. We easily see that T(0,0,1)=(0,0,5), so that 5 is an eigenvalue. Also $T(1,0,0)=(0,0,0)$ so 0 is an eigenvalue. Assume that $\lambda \neq 0$ and $T(z_1,z_2,z_3)=(2z_2,0,5z_3)=\lambda(z_1,z_2,z_3)$. From the assumption $\lambda \neq 0$ we get $z_2=0$, so the equation is of the form $(0,0,5z_3)=\lambda(z_1,0,z_3)$. Again we see that $z_1=0$ so we get the equation $(0,0,5z_3)=\lambda(0,0,z_3)$. Thus 5 is the only non-zero eigenvalue.

7. Notice that the range of T is the subspace $\{(x,\ldots,x) \in \mathbb{F}^n \mid x \in \mathbb{F}\}$ and has dimension 1. Thus $\dim \operatorname{range} T=1$, so $\dim \operatorname{null} T=n-1$. Assume that T has two distinct eigenvalues λ_1,λ_2 and assume that $\lambda_1 \neq 0 \neq \lambda_2$. Let v_1,v_2 be the corresponding eigenvectors, so by theorem 5.6 they are linearly independent. Then $v_1,v_2 \notin \operatorname{null} T$, but $\dim \operatorname{null} T=n-1$, so this is impossible. Hence T has at most one non-zero eigenvalue hence at most two eigenvalues.

   Because T is not injective we know that 0 is an eigenvalue. We also see that $T(1,\ldots,1)=n(1,\ldots,1)$, so n is another eigenvalue. By the previous paragraph, these are all eigenvalues of T.

8. Let $a \in \mathbb{F}$. Now we have $T(a,a^2,a^3,\ldots)=(a^2,a^3,\ldots)=a(a,a^2,\ldots)$, so every $a \in \mathbb{F}$ is an eigenvalue.

9. Assume that T has $k+2$ distinct eigenvalues $\lambda_1,\ldots,\lambda_{k+2}$ with corresponding eivenvectors $v_1,\ldots,v_{k+2}$. By theorem 5.6 these eigenvectors are linearly independent. Now $Tv_i=\lambda_i v_i$ and $\dim \operatorname{span}(Tv_1,\ldots,Tv_{k+2})=\dim \operatorname{span}(\lambda_1 v_1,\ldots,\lambda_{k+2}v_{k+2}) \geq k+1$ (it's $k+2$ if all λ_i are non-zero, otherwise $k+1$). This is a contradiction as $\dim \operatorname{range} T=k$ and $\operatorname{span}(\lambda_1 v_1,\ldots,\lambda_{k+2}v_{k+2}) \subset \operatorname{range} T$.

10. As $T=(T^{-1})^{-1}$ we only need to show this in one direction. If T is invertible, then 0 is not an eigenvalue. Now let λ be an eigenvalue of T and v the corresponding eigenvector. From $Tv=\lambda v$ we get that $T^{-1}\lambda v=\lambda T^{-1}v=v$, so that $T^{-1}v=\lambda^{-1}v$.

11. Let λ be an eigenvalue of TS and v the corresponding eigenvector. Then we get $STSv=S\lambda v=\lambda Sv$, so if $Sv \neq 0$, then it is is an eigenvector for the eigenvalue λ. If $Sv=0$, then $TSv=0$, so $\lambda=0$. As $Sv=0$ we know that S is not injective, so ST is not injective and it has eigenvalue 0. Thus if λ is an eigenvalue of TS, then it's an eigenvalue of ST. The other implication follows by symmetry.

12. Let $(v_1, \ldots, v_n)$ be a basis of V. By assumption $(Tv_1, \ldots, Tv_n) = (\lambda_1 v, \ldots, \lambda_n v_n)$. We need to show that $\lambda_i = \lambda_j$ for all i, j. Choose $i \neq j$, then by our assumption $v_i + v_j$ is an eigenvector, so $T(v_i + v_j) = \lambda_i v_i + \lambda_j v_j = \lambda(v_i + v_j) = \lambda v_i + \lambda v_j$. This means that $(\lambda_i - \lambda)v_i + (\lambda_j - \lambda)v_j = 0$. Because (v_i, v_j) is linearly independent we get that $\lambda_i - \lambda = 0 = \lambda_j - \lambda$ i.e. $\lambda_i = \lambda_j$.

13. Let $v_1 \in V$ and extend it to a basis $(v_1, \ldots, v_n)$ of V. Now $Tv_1 = a_1 v_1 + \ldots + a_n v_n$. Let U_i be the subspace generated by the vectors v_j, $j \neq i$. By our assumption each U_i is an invariant subspace. Let $Tv_1 = a_1 v_1 + \ldots + a_n v_n$. Now $v_1 \in U_i$ for $i > 1$. So let $j > 1$, then $Tv \in U_j$ imples $a_j = 0$. Thus $Tv_1 = a_1 v_1$. We see that v_1 was an eigenvector. The result now follows from the previous exercise.

14. Clearly $(STS^{-1})^n = ST^n S^{-1}$, so

$$\sum_{i=0}^{n} a_i (STS^{-1})^i = \sum_{i=0}^{n} a_i ST^i S^{-1} = S\left(\sum_{i=0}^{n} a_i T^i\right) S^{-1}.$$

15. Let λ be an eigenvalue of T and $v \in V$ a corresponding eigenvector. Let $p(x) = \sum_{i=0}^{n} a_i x^i$. Then we have

$$p(T)v = \sum_{i=0}^{n} a_i T^i v = \sum_{i=0}^{n} a_i \lambda^i v = p(\lambda)v.$$

Thus $p(\lambda)$ is an eigenvalue of $p(T)$. Then let a be an eigenvalue of $p(T)$ and $v \in V$ the corresponding eigenvector. Let $q(x) = p(x) - a$. By the fundamental theorem of algebra we can write $q(x) = c\prod_{i=1}^{n}(x - \lambda_i)$. Now $q(T)v = 0$ and $q(T) = c\prod_{i=1}^{n}(T - \lambda_i I)$. As $q(T)$ is non-injective we have that $T - \lambda_i I$ is non-injective for some i. Hence λ_i is an eigenvalue of T. Thus we get $0 = q(\lambda_i) = p(\lambda_i) - a$, so that $a = p(\lambda_i)$.

16. Let $T \in \mathcal{L}(\mathbb{R}^2)$ be the map $T(x, y) = (-y, x)$. On page 78 it was shown that it has no eigenvalue. However, $T^2(x, y) = T(-y, x) = (-x, -y)$, so -1 it has an eigenvalue.

17. By theorem 5.13 T has an upper-triangular matrix with respect to some basis $(v_1, \ldots, v_n)$. The claim now follows from proposition 5.12.

18. Let $T \in \mathcal{L}(\mathbb{F}^2)$ be the operator $T(a, b) = (b, a)$ with respect to the standard basis. Now $T^2 = I$, so T is clearly invertible. However, T has the matrix

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

19. Take the operator T in exercise 7. It is clearly not invertible, but has the matrix

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

20. By theorem 5.6 we can choose basis $(v_1, \ldots, v_n)$ for V where v_i is an eigenvector of T corresponding to the eigenvalue λ_i. Let λ be the eigenvalue of S corresponding to v_i. Then we get

$$STv_i = S\lambda_i v_i = \lambda_i S v_i = \lambda_i \lambda v_i = \lambda\lambda_i v_i = \lambda T v_i = T\lambda v_i = TSv_i,$$

so ST and TS agree on a basis of V. Hence they are equal.

21. Clearly 0 is an eigenvalue and the corresponding eigenvectors are $\operatorname{null} T = W \setminus \{0\}$. Now assume $\lambda \neq 0$ is an eigenvalue and $v = u + w$ is a corresponding eigenvector. Then $P_{U,W}(u + w) = u = \lambda u + \lambda w \Leftrightarrow (1 - \lambda)u = \lambda w$. Thus $\lambda w \in U$, so $\lambda w = 0$, but $\lambda \neq 0$, so we have $w = 0$ which implies $(1 - \lambda)u = 0$. Because v is an eigenvector we have $v = u + w = u \neq 0$, so that $1 - \lambda = 0$. Hence $\lambda = 1$. As we can choose freely our $u \in U$, so that it's non-zero we see that the eigenvectors corresponding to 1 are $u = U \setminus \{0\}$.

22. As $\dim V = \dim \operatorname{null} P + \dim \operatorname{range} P$, it's clearly sufficient to prove that $\operatorname{null} P \cap \operatorname{range} P = \{0\}$. Let $v \in \operatorname{null} P \cap \operatorname{range} P$. As $v \in \operatorname{range} P$, we can find a $u \in V$ such that $Pu = v$. Thus we have that $v = Pu = P^2 u = Pv$, but $v \in \operatorname{null} P$, so $Pv = 0$. Thus $v = 0$.

23. Let $T(a, b, c, d) = (-b, a, -d, c)$, then T is clearly injective, so 0 is not an eigenvalue. If $\lambda \neq 0$ is an eigenvalue and $\lambda(a, b, c, d) = (-b, a, -d, c)$. We clearly see that $a \neq 0 \Leftrightarrow b \neq 0$ and similarly with c, d. By symmetry we can assume that $a \neq 0$ (an eigenvector is non-zero). Then we have $\lambda a = -b$ and $\lambda b = a$. Substituting we get $\lambda^2 b = -b \Leftrightarrow (\lambda^2 + 1)b = 0$. As $b \neq 0$ we have $\lambda^2 + 1 = 0$, but this equation has no solutions in $\mathbb{R}$. Hence T has no eigenvalue.

24. If U is an invariant subspace of odd degree, then by theorem 5.26 $T_{|U}$ has an eigenvalue λ with eigenvector v. Then λ is an eigenvalue of T with eigenvector v against our assumption. Thus T has no subspace of odd degree.

6 Inner-Product Spaces

1. From the law of cosines we get $\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\|y\| \cos\theta$. Solving we get

$$\|x\|\|y\| \cos\theta = \frac{\|x\|^2 + \|y\|^2 - \|x - y\|^2}{2}.$$

Now let $x = (x_1, x2)$, $y = (y_1, y_2)$. A straight calculation shows that

$$\|x\|^2 + \|y\|^2 - \|x-y\|^2 = x_1^2 + x_2^2 + y_1^2 + y_2^2 - (x_1 - y_1)^2 - (x_2 - y_2)^2 = 2(x_1y_1 + x_2y_2),$$

so that

$$\|x\|\|y\| \cos\theta = \frac{\|x\|^2 + \|y\|^2 - \|x-y\|^2}{2} = \frac{2(x_1y_1 + x_2y_2)}{2} = \langle x, y \rangle .$$

2. If $\langle u, v \rangle = 0$, then $\langle u, av \rangle = 0$, so by the Pythagorean theorem $\|u\| \leq \|u\| + \|av\| = \|u+av\|$. Then assume that $\|u\| \leq \|u+av\|$ for all $a \in \mathbb{F}$. We have $\|u\|^2 \leq \|u+av\|^2 \Leftrightarrow \langle u, u \rangle \leq \langle u + av, u + av \rangle$. Thus we get

   $$\langle u, u \rangle \leq \langle u, u \rangle + \langle u, av \rangle + \langle av, u \rangle + \langle av, av \rangle ,$$

   so that

   $$-2\mathrm{Re}\,\overline{a} \langle u, v \rangle \leq |a|^2 \|v\|^2.$$

   Choose $a = -t \langle u, v \rangle$ with $t > 0$, so that

   $$2t \langle u, v \rangle^2 \leq t^2 \langle u, v \rangle^2 \|v\|^2 \Leftrightarrow 2 \langle u, v \rangle^2 \leq t \langle u, v \rangle^2 \|v\|^2.$$

   If $v = 0$, then clearly $\langle u, v \rangle = 0$. If not choose $t = 1/\|v\|^2$, so that we get

   $$2 \langle u, v \rangle^2 \leq \langle u, v \rangle^2 .$$

   Thus $\langle u, v \rangle = 0$.

3. Let $a = (a_1, \sqrt{2}a_2, \ldots, \sqrt{n}a_n) \in \mathbb{R}^n$ and $b = (b_1, b_2/\sqrt{2}, \ldots, b_n/\sqrt{n}) \in \mathbb{R}^n$. This equality is then simply $\langle a, b \rangle^2 \leq \|a\|^2\|b\|^2$ which follows directly from the Cauchy-Schwarz inequality.

4. From the parallelogram equality we have $\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$. Solving for $\|v\|$ we get $\|v\| = \sqrt{17}$.

5. Set e.g. $u = (1, 0)$, $v = (0, 1)$. Then $\|u\| = 1$, $\|v\| = 1$, $\|u + v\| = 2$, $\|u - v\| = 2$. Assuming that the norm is induced by an inner-product, we would have by the parallelogram inequality

   $$8 = 2^2 + 2^2 = 2(1^2 + 1^2) = 4,$$

   which is clearly false.

6. Just use $\|u\| = \langle u, u \rangle$ and simplify.

7. See previous exercise.

8. This exercise is a lot trickier than it might seem. I'll prove it for $\mathbb{R}$, the proof for $\mathbb{C}$ is almost identical except for the calculations that are longer and more tedious. All norms on a finite-dimensional real vector space are equivalent. This gives us

$$\lim_{n\to\infty} \|r_n x + y\| = \|\lambda x + y\|$$

when $r_n \to \lambda$. This probably doesn't make any sense, so check a book on topology or functional analysis or just assume the result.

Define $\langle u, v\rangle$ by

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4}.$$

Trivially we have $\|u\|^2 = \langle u, u\rangle$, so positivity definiteness follows. Now

$$\begin{aligned}
4(\langle u+v, w\rangle - \langle u, w\rangle - \langle v, w\rangle) &= \|u+v+w\|^2 - \|u+v-w\|^2 \\
&\quad -(\|u+w\|^2 - \|u-w\|^2) \\
&\quad -(\|v+w\|^2 - \|v-w\|^2) \\
&= \|u+v+w\|^2 - \|u+v-w\|^2 \\
&\quad +(\|u-w\|^2 + \|v-w\|^2) \\
&\quad -(\|u+w\|^2 + \|v+w\|^2)
\end{aligned}$$

We can apply the parallelogram equality to the two parenthesis getting

$$\begin{aligned}
4(\langle u+v, w\rangle - \langle u, w\rangle - \langle v, w\rangle) &= \|u+v+w\|^2 - \|u+v-w\|^2 \\
&+ \frac{1}{2}(\|u+w-2w\|^2 + |u-v|^2) \\
&- \frac{1}{2}(\|u+v+2w\|^2 + \|u-v\|^2) \\
&= \|u+v+w\|^2 - \|u+v-w\|^2 \\
&+ \frac{1}{2}\|u+w-2w\|^2 - \frac{1}{2}\|u+v+2w\|^2 \\
&= (\|u+v+w\|^2 + \|w\|^2) + \frac{1}{2}\|u+v-2w\|^2 \\
&- (\|u+v-w\|^2 + \|w\|^2) - \frac{1}{2}\|u+v+2w\|^2
\end{aligned}$$

Applying the parallelogram equality to the two parenthesis we and simplifying we are left with 0. Hence we get additivity in the first slot. It's easy to see that $\langle -u, v\rangle = -\langle u, v\rangle$, so we get $\langle nu, v\rangle = n\langle u, v\rangle$ for all $n \in \mathbb{Z}$. Now we get

$$\langle u, v\rangle = \langle nu/n, v\rangle = n\langle u/n, v\rangle,$$

so that $\langle u/n, v\rangle = 1/n\,\langle u, v\rangle$. Thus we have $\langle nu/m, v\rangle = n/m\,\langle u, v\rangle$. It follows that we have homogeneity in the first slot when the scalar is rational. Now let $\lambda \in \mathbb{R}$ and choose a sequence (r_n) of rational numbers such that $r_n \to \lambda$. This gives us

$$\begin{aligned}
\lambda\,\langle u, v\rangle &= \lim_{n\to\infty} r_n\,\langle u, v\rangle = \lim_{n\to\infty}\langle r_n u, v\rangle \\
&= \lim_{n\to\infty}\frac{1}{4}(\|r_n u + v\|^2 - \|r_n u - v\|^2) \\
&= \frac{1}{4}(\|\lambda u + v\|^2 - \|\lambda u - v\|^2) \\
&= \langle \lambda u, v\rangle
\end{aligned}$$

Thus we have homogeneity in the first slot. We trivially also have symmetry, so we have an inner-product. The proof for $\mathbb{F} = \mathbb{C}$ can be done by defining the inner-product from the complex polarizing identity. And by using the identities

$$\langle u, v\rangle_0 = \frac{1}{4}(\|u + v\|^2 - \|u - v\|^2) \Rightarrow \langle u, v\rangle = \langle u, v\rangle_0 + \langle u, iv\rangle_0\, i.$$

and using the properties just proved for $\langle \cdot, \cdot\rangle_0$.

9. This is really an exercise in calculus. We have integrals of the types

$$\begin{aligned}
\langle \sin nx, \sin mx\rangle &= \int_{-\pi}^{\pi} \sin nx \sin mx dx \\
\langle \cos nx, \cos mx\rangle &= \int_{-\pi}^{\pi} \cos nx \cos mx dx \\
\langle \sin nx, \cos mx\rangle &= \int_{-\pi}^{\pi} \sin nx \cos mx dx
\end{aligned}$$

and they can be evaluated using the trigonometric identities

$$\begin{aligned}
\sin nx \sin mx &= \frac{\cos((n-m)x) - \cos((n+m)x)}{2} \\
\cos nx \cos mx &= \frac{\cos((n-m)x) + \cos((n+m)x)}{2} \\
\cos nx \sin mx &= \frac{\sin((n-m)x) - \sin((n+m)x)}{2}.
\end{aligned}$$

10. $e_1 = 1$, then $e_2 = \frac{x - \langle x, 1\rangle 1}{\|x - \langle x, 1\rangle 1\|}$. Where

$$\langle x, 1\rangle = \int_0^1 x = 1/2,$$

so that

$$\|x-1/2\| = \langle x-1/2, x-1/2\rangle = \sqrt{\int_0^1 (x-1/2)^2 dx} = \sqrt{1/12},$$

which gives $e_2 = \sqrt{12}(x-1/2) = \sqrt{3}(2x-1)$. Then to continue pencil pushing we have

$$\left\langle x^2, 1\right\rangle = 1/3, \left\langle x^2, \sqrt{3}(2x-1)\right\rangle = \sqrt{3}/6,$$

so that $x^2 - \left\langle x^2, 1\right\rangle 1 - \left\langle x^2, \sqrt{3}(2x-1)\right\rangle \sqrt{3}(2x-1) = x^2 - 1/3 - 1/2(2x-1)$. Now

$$\|x^2 - 1/3 - 1/2(2x-1)\| = 1/(6\sqrt{5})$$

giving $e_3 = \sqrt{5}(6x^2-6x+1)$. The orthonormal basis is thus $(1, \sqrt{3}(2x-1), \sqrt{5}(6x^2 - 6x+1)$.

11. Let $(v_1, \ldots, v_n)$ be a linearly dependent list. We can assume that $(v_1, \ldots, v_{n-1})$ is linearly independent and $v_n \in \operatorname{span}(v_1, \ldots, v_n)$. Let P denote the projection on the subspace spanned by $(v_1, \ldots, v_{n-1})$. Then calculating e_n with the Gram-Schmidt algorithm gives us

$$e_n = \frac{v_n - P(v_n)}{\|v_n - Pv_n\|},$$

but $v_n = Pv_n$ by our assumption, so $e_n = 0$. Thus the resulting list will simply contain zero vectors. It's easy to see that we can extend the algorithm to work for linearly dependent lists by tossing away resulting zero vectors.

12. Let $(e_1, \ldots, e_{n-1})$ be an orthogonal list of vectors. Assume that $(e_1, \ldots, e_n)$ and $(e_1, \ldots, e_{n-1}, e_n')$ are orthogonal and span the same subspace. Then we can write $e_n' = a_1e_1 + \ldots + a_ne_n$. Now we have $\langle e_n', e_i\rangle = 0$ for all $i < n$, so that $a_i = 0$ for $i < n$. Thus we have $e_n' = a_ne_n$ and from $\|e_n'\| = 1$ we have $a_n = \pm 1$.

    Let $(e_1, \ldots, e_n)$ be the orthonormal base produced from $(v_1, \ldots, v_n)$ by Gram-Schmidt. Then if $(e_1', \ldots, e_n')$ satisfies the hypothesis from the problem we have by the previous paragraph that $e_i' = \pm e_i$. Thus we have 2^n possible such orthonormal lists.

13. Extend $(e_1, \ldots, e_m)$ to a basis $(e_1, \ldots, e_n)$ of V. By theorem 6.17

$$v = \langle v, e_1\rangle e_1 + \ldots + \langle v, e_n\rangle e_n,$$

so that

$$\begin{aligned} \|v\| &= \| \langle v, e_1\rangle e_1 + \ldots + \langle v, e_m\rangle e_n\| \\ &= \| \langle v, e_1\rangle e_1\| + \ldots + \| \langle v, e_m\rangle e_n\| \\ &= | \langle v, e_1\rangle | + \ldots + | \langle v, e_n\rangle |. \end{aligned}$$

Thus we have $\|v\| = | \langle v, e_1\rangle | + \ldots + | \langle v, e_m\rangle |$ if and only if $\langle v, e_i\rangle = 0$ for $i > m$ i.e. $v \in \operatorname{span}(e_1, \ldots, e_m)$.

14. It's easy to see that the differentiation operator has a upper-triangular matrix in the orthonormal basis calculated in exercise 10.

15. This follows directly from $V = U \oplus U^{\perp}$.

16. This follows directly from the previous exercise.

17. From exercise 5.21 we have that $V = \operatorname{range} P \oplus \operatorname{null} P$. Let $U := \operatorname{range} P$, then by assumption $\operatorname{null} P \subset U^{\perp}$. From exercise 15, we have $\dim U = \dim V - \dim U = \dim \operatorname{null} P$, so that $\operatorname{null} P = U^{\perp}$. An arbitrary $v \in V$ can be written as $v = Pv + (v - Pv)$. From $P^2 = P$ we have that $P(v - Pv) = 0$, so $v - Pv \in \operatorname{null} P = U^{\perp}$. Hence the decomposition $v = Pv + (v - Pv)$ is the unique decomposition in $U \oplus U^{\perp}$. Now $Pv = P(Pv + (v - Pv)) = Pv$, so that P is the identity on U. By definition $P = P_U$.

18. Let $u \in \operatorname{range} P$, then $u = Pv$ for some $v \in V$ hence $Pu = P^2 v = Pv = u$, so P is the identity on $\operatorname{range} P$. Let $w \in \operatorname{null} P$. Then for $a \in \mathbb{F}$

$$\|u\|^2 = \|P(u + aw)\|^2 \leq \|u + aw\|^2.$$

By exercise 2 we have $\langle u, w \rangle = 0$. Thus $\operatorname{null} P \subset (\operatorname{range} P)^{\perp}$ and from the dimension equality $\operatorname{null} P = (\operatorname{range} P)^{\perp}$. Hence the claim follows.

19. If $TP_U = P_U T P_U$ clearly U is invariant. Now assume that U is invariant. Then for $u \in U$ we have $P_U T P_U u = P_U T u = T u = T P_U u$.

20. Let $u \in U$, then we have $Tu = TP_U u = P_U T u \in U$. Thus U is invariant. Then let $w \in U^{\perp}$. Now we can write $Tw = u + u'$ where $u \in U$ and $u' \in U^{\perp}$. Now $P_U T w = u$, but $u = P_U T w = T P_U w = T0 = 0$, so $Tw \in U^{\perp}$. Thus $U^{\perp}$ is also invariant.

21. First we need to find an orthonormal basis for U. With Gram-Schmidt we get $e_1 = (1/\sqrt{2}, 1/\sqrt{2}, 0, 0)$, $e_2 = (0, 0, 1/\sqrt{5}, 2/\sqrt{5})$. Let $U = \operatorname{span}(e_1, e_2)$. Then we have

$$u = P_U(1,2,3,4) = \langle (1,2,3,4), e_1 \rangle e_1 + \langle (1,2,3,4), e_2 \rangle e_2 = (3/2, 3/2, 11/5, 22/5).$$

22. If $p(0) = 0$ and $p'(0) = 0$, then $p(x) = ax^2 + bx^3$. Thus we want to find the projection of $2 + 3x$ to the subspace $U := \operatorname{span}(x^2, x^3)$. With Gram-Schmidt we get the orthonormal basis $(\sqrt{3}x^2, \sqrt{420/11}(x^3 - \frac{1}{2}x^2))$. We get

$$P_U(2+3x) = \left\langle 2 + 3x, \sqrt{3}x^2 \right\rangle \sqrt{3}x^2 + \left\langle 2 + 3x, \sqrt{420/11}(x^3 - \frac{1}{2}x^2) \right\rangle \sqrt{420/11}(x^3 - \frac{1}{2}x^2).$$

Here

$$\left\langle 2 + 3x, \sqrt{3}x^2 \right\rangle = \frac{17}{12}\sqrt{3}, \quad \left\langle 2 + 3x, \sqrt{420/11}(x^3 - \frac{1}{2}x^2) \right\rangle = \frac{47}{660}\sqrt{1155},$$

so that

$$P_U(2 + 3x) = -\frac{71}{22}x^2 + \frac{329}{22}x^3.$$

23. There's is nothing special with this exercise, compared to the previous two, except that it takes ages to calculate.

24. We see that the map $T : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ defined by $p \mapsto p(1/2)$ is linear. From exercise 10 we get that $(1, \sqrt{3}(2x-1), \sqrt{5}(6x^2 - 6x + 1)$ is an orthonormal basis for $\mathcal{P}_2(\mathbb{R})$. From theorem 6.45 we see that

$$\begin{aligned} q(x) &= 1 + \sqrt{3}(2 \cdot 1/2 - 1)\sqrt{3}(2x-1) + \sqrt{5}(6 \cdot (1/2)^2 - 6 \cdot 1/2 + 1)\sqrt{5}(6x^2 - 6x + 1) \\ &= -3/2 + 15x - 15x^2. \end{aligned}$$

25. The map $T : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ defined by $p \mapsto \int_0^1 p(x)\cos(\pi x)dx$ is clearly linear. Again we have the orthonormal basis $(1, \sqrt{3}(2x-1), \sqrt{5}(6x^2 - 6x + 1)$, so that

$$\begin{aligned} q(x) &= \int_0^1 \cos(\pi x)dx + \int_0^1 \sqrt{3}(2x-1)\cos(\pi x)dx \times \sqrt{3}(2x-1) \\ &+ \int_0^1 \sqrt{5}(6x^2 - 6x + 1)\cos(\pi x)dx \times \sqrt{5}(6x^2 - 6x + 1) \\ &= 0 - \frac{4\sqrt{3}}{\pi^2}\sqrt{3}(2x-1) + 0 = -\frac{12}{\pi^2}(2x-1). \end{aligned}$$

26. Choose an orthonormal basis $(e_1, \ldots, e_n)$ for V and let $\mathbb{F}$ have the usual basis. Then by proposition 6.47 we have

$$\mathcal{M}(T, (e_1, \ldots, e_n)) = [\langle e_1, v\rangle \cdots \langle e_n, v\rangle],$$

so that

$$\mathcal{M}(T^*, (e_1, \ldots, e_n)) = \begin{bmatrix} \overline{\langle e_1, v\rangle} \\ \vdots \\ \overline{\langle e_n, v\rangle} \end{bmatrix}$$

and finally $T^*a = (a\overline{\langle e_1, v\rangle}, \ldots, a\overline{\langle e_n, v\rangle})$.

27. with the usual basis for $\mathbb{F}^n$ we get

$$\mathcal{M}(T) = \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix}$$

so by proposition 6.47

$$\mathcal{M}(T^*) = \begin{bmatrix} 0 & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & 0 \end{bmatrix}.$$

Clearly $T^*(z_1, \dots, z_n) = (z_2, \dots, z_n, 0)$.

28. By additivity and conjugate homogeneity we have $(T - \lambda I)^* = T^* - \overline{\lambda} I$. From exercise 31 we see that $T^* - \overline{\lambda} I$ is non-injective if and only if $T - \lambda I$ is non-injective. That is $\overline{\lambda}$ is an eigenvalue of T^* if and only if λ is an eigenvalue of T.

29. As $(U^\perp)^\perp = U$ and $(T^*)^* = T$ it's enough to show only one of the implications. Let U be invariant under T and choose $w \in U^\perp$. Now we can write $T^* w = u + v$, where $u \in U$ and $v \in U^\perp$, so that

$$0 = \langle Tu, w \rangle = \langle u, T^* w \rangle = \langle u, u + v \rangle = \langle u, u \rangle = \|u\|^2.$$

Thus $u = 0$ which completes the proof.

30. As $(T^*)^* = T$ it's sufficient to prove the first part. Assume that T is injective, then by exercise 31 we have

$$\dim V = \dim \operatorname{range} T = \dim \operatorname{range} T^*,$$

but $\operatorname{range} T$ is a subspace of V, so that $\operatorname{range} T = V$.

31. From proposition 6.46 and exercise 15 we get

$$\begin{aligned} \dim \operatorname{null} T &= \dim(\operatorname{range} T)^\perp = \dim W - \dim \operatorname{range} T \\ &= \dim \operatorname{null} T + \dim W - \dim V. \end{aligned}$$

For the second part we have

$$\begin{aligned} \dim \operatorname{range} T^* &= \dim W - \dim \operatorname{null} T^* \\ &= \dim V - \dim \operatorname{null} T \\ &= \dim \operatorname{range} T, \end{aligned}$$

where the second equality follows from the first part.

32. Let T be the operator induced by A. The columns are the images of the basis vectors under T, so the generate $\operatorname{range} T$. Hence the dimension of the span of the column vectors equal $\dim \operatorname{range} T$. By proposition 6.47 the span of the row vectors equals $\operatorname{range} T^*$. By the previous exercise $\dim \operatorname{range} T = \dim \operatorname{range} T^*$, so the claim follows.

7 Operators on Inner-Product Spaces

1. For the first part, let $p(x) = x$ and $q(x) = 1$. Then clearly $1/2 = \langle Tp, q \rangle \neq \langle p, Tq \rangle = \langle p, 0 \rangle = 0$. For the second part it's not a contradiction since the basis is not orthogonal.

2. Choose the standard basis for $\mathbb{R}^2$. Let T and S the the operators defined by the matrices
$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$
Then T and S as clearly self-adjoint, but TS has the matrix
$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$
so TS is not self-adjoint.

3. If T and S are self-adjoint, then $(aT+bS)^* = (aT)^*+(bS)^* = aT+bS$ for any $a, b \in \mathbb{R}$, so self-adjoint operators form a subspace. The identity operator I is self-adjoint, but clearly $(iI)^* = -iI$, so iI is not self-adjoint.

4. Assume that P is self-adjoint. Then from proposition 6.46 we have $\operatorname{null} P = \operatorname{null} P^* = (\operatorname{range} P)^{\perp}$. Now P is clearly a projection to $\operatorname{range} P$. Then assume that P is a projection. Let $(u_1, \ldots, u_n)$ an orthogonal basis for $\operatorname{range} P$ and extend it to a basis of V. Then clearly P has a diagonal matrix with respect to this basis, so P is self-adjoint.

5. Choose the operators corresponding to the matrices
$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \ B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$
Then we easily see that $A^*A = AA^*$ and $B^*B = BB^*$. However, $A+B$ doesn't define a normal operator which is easy to check.

6. From proposition 7.6 we get $\operatorname{null} T = \operatorname{null} T^*$. It follows from proposition 6.46 that
$$\operatorname{range} T = (\operatorname{null} T^*)^{\perp} = (\operatorname{null} T)^{\perp} = \operatorname{range} T^*.$$

7. Clearly $\operatorname{null} T \subset \operatorname{null} T^k$. Let $v \in \operatorname{null} T^k$. Then we have
$$\left\langle T^*T^{k-1}v, T^*T^{k-1}v \right\rangle = \left\langle T^*T^k v, T^{k-1}v \right\rangle = 0,$$
so that $T^*T^{k-1}v = 0$. Now
$$\left\langle T^{k-1}v, T^{k-1}v \right\rangle = \left\langle T^*T^{k-1}v, T^{k-2}v \right\rangle = 0,$$
so that $v \in \operatorname{null} T^{k-1}$. Thus $\operatorname{null} T^k \subset \operatorname{null} T^{k-1}$ and continuing we get $\operatorname{null} T^k \subset \operatorname{null} T$.

   Now let $u \in \operatorname{range} T^k$, then we can find a $v \in V$ such that $u = T^k v = T(T^{k-1}v)$, so that $\operatorname{range} T^k \subset \operatorname{range} T$. From the first part we get $\dim \operatorname{range} T^k = \dim \operatorname{range} T$, so $\operatorname{range} T^k = \operatorname{range} T$.

8. The vectors $u = (1, 2, 3)$ and $v = (2, 5, 7)$ are both eigenvectors corresponding to different eigenvalues. A self-adjoint operator is normal, so by corollary 7.8 if T is normal that would imply orthogonality of u and v. Clearly $\langle u, v\rangle \neq 0$, so T can't be even normal much less self-adjoint.

9. If T is normal, we can choose a basis of V consisting of eigenvectors of T. Let A be the matrix of T corresponding to the basis. Now T is self-adjoint if and only if the conjugate transpose of A equals A that is the eigenvalues of T are real.

10. For any $v \in V$ we get $0 = T^9 v - T^8 v = T^8(Tv - v)$. Thus $Tv - v \in \operatorname{null} T^8 = \operatorname{null} T$ by the normality of T. Hence $T(Tv - v) = T^2 v - Tv = 0$, so $T^2 = T$. By the spectral theorem we can choose a basis of eigenvectors for T such that T has a diagonal matrix with $\lambda_1, \ldots, \lambda_n$ on the diagonal. Now T^2 has the diagonal matrix with $\lambda_1^2, \ldots, \lambda_n^2$ on the diagonal and from $T^2 = T$ we must have $\lambda_i^2 = \lambda_i$ for all i. Hence $\lambda_i = 0$ or $\lambda_i = 1$, so the matrix of T equals its conjugate transpose. Hence T is self-adjoint.

11. By the spectral theorem we can choose a basis of T such that the matrix of T corresponding to the basis is a diagonal matrix. Let $\lambda_1, \ldots, \lambda_n$ be the diagonal elements. Let S be the operator corresponding to the diagonal matrix having the elements $\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}$ on the diagonal. Clearly $S^2 = T$.

12. Let T the operator corresponding to the matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Then
$$T^2 + I = 0.$$

13. By the spectral theorem we can choose a basis consisting of eigenvectors of T. Then T has a diagonal matrix with respect to the basis. Let $\lambda_1, \ldots, \lambda_n$ be the diagonal elements. Let S be the operator corresponding to the diagonal matrix having $\sqrt[3]{\lambda_1}, \ldots, \sqrt[3]{\lambda_n}$ on the diagonal. Then clearly $S^3 = T$.

14. By the spectral theorem we can choose a basis $(v_1, \ldots, v_n)$ consisting of eigenvectors of T. Let $v \in V$ be such that $\|v\| = 1$. If $v = a_1 v_1 + \ldots + a_n v_n$ this implies that $\sum_{i=1}^n |a_i| = 1$. Assume that $\|Tv - \lambda v\| < \varepsilon$. If $|\lambda_i - \lambda| \geq \varepsilon$ for all i, then
$$\begin{aligned} \|Tv - \lambda v\| &= \|\sum_{i=1}^n (a_i T v_i - \lambda v_i)\| = \|\sum_{i=1}^n (a_i \lambda_i v_i - \lambda v_i)\| \\ &= \|\sum_{i=1}^n a_i(\lambda_i - \lambda) v_i\| = \sum_{i=1}^n |a_i||\lambda_i - \lambda|\|v_1\| \\ &= \sum_{i=1}^n |a_i||\lambda_i - \lambda| \geq \varepsilon \sum_{i=1}^n |a_i| = \varepsilon, \end{aligned}$$
which is a contradiction. Thus we can find an eigenvalue λ_i such that $|\lambda - \lambda_i| < \varepsilon$.

15. If such an inner-product exists we have a basis consisting of eigenvectors by the spectral theorem. Assume then that $(v_1, \ldots, v_n)$ is basis such that the matrix of T is a matrix consisting of eigenvectors of T. Define an inner-product by

$$\langle v_i, v_j \rangle = \begin{cases} 0, i \neq j \\ 1, i = j \end{cases} .$$

Extend this to an inner-product by bilinearity and homogeneity. Then we have an inner-product and clearly T is self-adjoint as $(v_1, \ldots, v_n)$ is an orthonormal basis of U.

16. Let $T \in \mathcal{L}(\mathbb{R}^2)$ be the operator $T(a, b) = (a, a + b)$. Then $\operatorname{span}((1, 0))$ is invariant under T, but it's orthogonal complement $\operatorname{span}((0, 1))$ is clearly not.

17. Let T, S be two positive operators. Then they are self-adjoint and $(S + T)^* = S^* + T^* = S + T$, so $S + T$ is self-adjoint. Also $\langle (S + T)v, v\rangle = \langle Sv, v\rangle + \langle Tv, v\rangle \geq 0$. Thus, $S + T$ is positive.

18. Clearly T^k is self-adjoint for every positive integer. For $k = 2$ we have $\left\langle T^2 v, v\right\rangle = \langle Tv, Tv\rangle \geq 0$. Assume that the result is true for all positive $k < n$ and $n \geq 2$. Then

$$\langle T^n v, v\rangle = \left\langle T^{n-2} Tv, Tv\right\rangle \geq 0,$$

by hypothesis and self-adjointness. Hence the result follows by induction.

19. Clearly $\langle Tv, v\rangle > 0$ for all $v \in V \setminus \{0\}$ implies that T is injective hence invertible. So assume that T is invertible. Since T is self-adjoint we can choose an orthonormal basis $(v_1, \ldots, v_n)$ of eigenvectors such that T has a diagonal matrix with respect to the basis. Let the elements on the diagonal be $\lambda_1, \ldots, \lambda_n$ and by assumption $\lambda_i > 0$ for all i. Let $v = a_1 v_1 + \ldots + a_n v_n \in V \setminus \{0\}$, so that

$$\langle Tv, v\rangle = |a_1|^2 \lambda_1 + \ldots + |a_n|^2 \lambda_n > 0.$$

20. $\begin{bmatrix} \sin\theta & \cos\theta \\ \cos\theta & -\sin\theta \end{bmatrix}$ is a square root for every $\theta \in \mathbb{R}$, which shows that it has infinitely many square roots.

21. Let $S \in \mathcal{L}(\mathbb{R}^2)$ be defined by $S(a, b) = (a + b, 0)$. Then $\|S(1, 0)\| = \|(1, 0)\| = 1$ and $\|S(0, 1)\| = 1$, but S is clearly not an isometry.

22. $\mathbb{R}^3$ is an odd-dimensional real vector space. Hence S has an eigenvalue λ and a corresponding eigenvector v. Since $\|Sv\| = |\lambda|\|v\| = \|v\|$ we have $\lambda^2 = 1$. Hence $S^2 v = \lambda^2 v = v$.

23. The matrices corresponding to T and T^* are

$$\mathcal{M}(T) = \begin{bmatrix} 0 & 0 & 1 \\ 2 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}, \ \mathcal{M}(T^*) = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{bmatrix}.$$

Thus we get

$$\mathcal{M}(T^*T) = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

so that T^*T is the operator $(z_1, z_2, z_3) \mapsto (4z_1, 9z_2, z_3)$. Hence $\sqrt{T^*T}(z_1, z_2, z_3) = (2z_1, 3z_2, z_3)$. Now we just need to permute the indices which corresponds to the isometry $S(z_1, z_2, z_3) = (z_3, z_1, z_2)$.

24. Clearly T^*T is positive, so by proposition 7.26 it has a unique positive square root. Thus it's sufficient to show that $R^2 = T^*T$. Now we have $T^* = (SR)^* = R^*S^* = RS^*$ by self-adjointness of R. Thus $R^2 = RIR = RS^*SR = T^*T$.

25. Assume that T is invertible. Then $\sqrt{T^*T}$ must be invertible (polar decomposition). Hence $S = T\sqrt{T^*T}^{-1}$, so S is uniquely determined. Now assume that T is not invertible. Then range $\sqrt{T^*T}$ is not invertible (hence not surjective).

   Now assume that T is not invertible, so that $\sqrt{T^*T}$ is not invertible. By the spectral theorem we can choose a basis $(v_1, \ldots, v_n)$ of eigenvectors of $\sqrt{T^*T}$ and we can assume that v_1 corresponds to the eigenvalue 0. Now let $T = S\sqrt{T^*T}$. Define U by $Uv_1 = -Sv_1$, $Uv_i = Sv_i$, $i > 1$. Clearly $U \neq S$ and $T = U\sqrt{T^*T}$. Now choose $u, v \in V$. Then it's easy to verify that $\langle Uu, Uv \rangle = \langle Su, Sv \rangle = \langle u, v \rangle$, so U is an isometry.

26. Choose a basis of eigenvectors for T such that T has a diagonal matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ on the diagonal. Now $T^*T = T^2$ corresponds to the diagonal matrix with $\lambda_1^2, \ldots, \lambda_n^2$ on the diagonal and the square root $\sqrt{T^*T}$ corresponds to the diagonal matrix having $|\lambda_1|, \ldots, |\lambda_n|$ on the diagonal. These are the singular values, so the claim follows.

27. Let $T \in \mathcal{L}(\mathbb{R}^2)$ be the map $T(a, b) = (0, a)$. Then from the matrix representation we see that $T^*T(a, 0) = (a, 0)$, hence $\sqrt{T^*T} = T^*T$ and 1 is clearly a singular value. However, $T^2 = 0$, so $\sqrt{(T^2)^*T^2} = 0$, so $1^2 = 1$ is not a singular value of T^2.

28. The composition of two bijections is a bijection. If T=$S\sqrt{T^*T}$, then since S is an isometry we have that T is bijective hence invertible if and only if $\sqrt{T^*T}$ is bijective. But $\sqrt{T^*T}$ is injective hence bijective if and only if 0 is not an eigenvalue. Thus T is bijective if and only if 0 is not a singular value.

29. From the polar decomposition theorem we see that $\dim \operatorname{range} T = \dim \operatorname{range} \sqrt{T^*T}$. Since $\sqrt{T^*T}$ is self-adjoint we can choose a basis of eigenvectors of $\sqrt{T^*T}$ such that the matrix of $\sqrt{T^*T}$ is a diagonal matrix. Clearly $\dim \operatorname{range} \sqrt{T^*T}$ equals the number of non-zero elements on the diagonal i.e. the number of non-zero singular values.

30. If S is an isometry, then clearly $\sqrt{S^*S} = I$, so all singular values are 1. Now assume that all singular values are 1. Then $\sqrt{S^*S}$ is a self-adjoint (positive) operator with all eigenvalues equal to 1. By the spectral theorem we can choose a basis of V such that $\sqrt{S^*S}$ has a diagonal matrix. As all eigenvalues are one, this means that the matrix of $\sqrt{S^*S}$ is the identity matrix. Thus $\sqrt{S^*S} = I$, so S is an isometry.

31. Let $T_1 = S_1\sqrt{T_1^*T_1}$ and $T_2 = S_2\sqrt{T_2^*T_2}$. Assume that T_1 and T_2 have the same singular values $s_1, \ldots, s_n$. Then we can choose bases of eigenvectors of $\sqrt{T_1^*T_1}$ and $\sqrt{T_2^*T_2}$ such that they have the same matrix. Let these bases be $(v_1, \ldots, v_n)$ and $(w_1, \ldots, w_n)$. Let S the operator defined by $S(w_i) = v_i$, then clearly S is an isometry and $\sqrt{T_2^*T_2} = S^{-1}\sqrt{T_1^*T_1}S$. Thus we get $T_2 = S_2\sqrt{T_2^*T_2} = S_2S^{-1}\sqrt{T_1^*T_1}S$. Writing $S_3 = S_2S^{-1}S_1^{-1}$, which is clearly an isometry, we get

$$T_2 = S_3S_1\sqrt{T_1^*T_1}S = S_3T_1S.$$

Now let $T_2 = S_1T_1S_2$. Then we have $T_2^*T_2 = (S_1T_1S_2)^*(S_1T_1S_2) = S_1^*T_1^*T_1S_1 = S_1^{-1}T_1^*T_1S_1$. Let λ be an eigenvalue of $T_1^*T_1$ and v a corresponding eigenvector. Because S_1 is bijective we have a $u \in V$ such that $v = S_1u$. This gives us

$$T_2^*T_2u = S_1^{-1}T_1^*T_1S_1u = S_1^{-1}T_1^*T_1v = S_1^{-1}\lambda v = \lambda u,$$

so that λ is an eigenvalue of $T_2^*T_2$. Let $(v_1, \ldots, v_n)$ be an orthonormal basis of eigenvectors of $T_1^*T_1$, then $(S_1^{-1}v_1, \ldots, S^{-1}v_n)$ is an orthonormal basis for V of eigenvectors of $T_2^*T_2$. Hence $T_2^*T_2$ and $T_1^*T_1$ have the same eigenvalues. The singular values are simply the positive square roots of these eigenvalues, so the claim follows.

32. For the first part, denote by S the map defined by the formula. Set

$$A := \mathcal{M}(S, (f_1, \ldots, f_n), (e_1, \ldots, e_n)),\ B := \mathcal{M}(T, (e_1, \ldots, e_n), (f_1, \ldots, f_n)).$$

Clearly $A = B$ and is a diagonal matrix with $s_1, \ldots, s_n$ on the diagonal. All the s_i are positive real numberse, so that $B^* = B = A$. It follows from proposition 6.47 that $S = T^*$.

For the second part observe that a linear map S defined by the formula maps f_i to $s_i^{-1}e_i$ which is mapped by T to f_i. Thus $TS = I$. The claim now follows from exercise 3.23.

33. By definition $\sqrt{T^*T}$ is positive. Thus all singular values of T are positive. From the singular value decomposition theorem we can choose bases of V such that

$$\|Tv\| = \| \sum_{i=1}^{n} s_i \langle v, e_i \rangle f_i \| = \sum_{i=1}^{n} s_i | \langle v, e_i \rangle |$$

since all s_i are positive. Now clearly

$$\hat{s}\|v\| = \hat{s} \sum_{i=1}^{n} | \langle v, e_i \rangle | \leq \sum_{i=1}^{n} s_i | \langle v, e_i \rangle | = \|Tv\|$$

and similarly for $s\|v\|$.

34. From the triangle equality and the previous exercise we get

$$\|(T' + T'')v\| \leq \|T'v\| + \|T''v\| \leq (s' + s'')\|v\|.$$

Now let v be the eigenvector of $\sqrt{(T' + T'')^*(T' + T'')}$ corresponding to the singular value s. Let $T + T'' = S\sqrt{(T' + T'')^*(T' + T'')}$, so that

$$\|(T' + T'')v\| = \|S\sqrt{(T' + T'')^*(T' + T'')}v\| = \|Ssv\| = s\|v\| \leq (s' + s'')\|v\|,$$

so that $s \leq s' + s''$.

8 Operators on Complex Vector Spaces

1. T is not injective, so 0 is an eigenvalue of T. We see that $T^2 = 0$, so that any $v \in V$ is a generalized eigenvector corresponding to the eigenvalue 0.

2. On page 78 it's shown that $\pm i$ are the eigenvalues of T. We have $\dim \operatorname{null}(T - iI)^2 + \dim \operatorname{null}(T + iI)^2 = \dim \mathbb{C}^2 = 2$, so that $\dim \operatorname{null}(T - iI)^2 = \dim \operatorname{null}(T + iI)^2 = 1$. Because $(1, -i)$ is an eigenvector corresponding to the eigenvalue i and $(1, i)$ is an eigenvector corresponding to the eigenvalue $-i$. The set of all generalized eigenvectors are simply the spans of these corresponding eigenvectors.

3. Let $\sum_{i=0}^{m-1} a_i T^i v = 0$, then $0 = T^{m-1}(\sum_{i=0}^{m-1} a_i T^i v = a_0 T^{m-1} v$, so that $a_0 = 0$. Applying repeatedly T^{m-i} for $i = 2, \ldots$ we see that $a_i = 0$ for all i. Hence $(v, \ldots, T^{m-1}v)$ is linearly independent.

4. We see that $T^3 = 0$, but $T^2 \neq 0$. Assume that S is a square root of T, then S is nilpotent, so by corollary 8.8 we have $S^{\dim V} = S^3 = 0$, so that $0 = S^4 = T^2$. Contradiction.

5. If ST is nilpotent, then $\exists n \in \mathbb{N}$ such that $(ST)^n = 0$, so $(TS)^{n+1} = T(ST)^n S = 0$.

6. Assume that $\lambda \neq 0$ is an eigenvalue of N with corresponding eigenvector v. Then $N^{\dim V} v = \lambda^{\dim V} v \neq 0$. This contradicts corollary 8.8.

7. By lemma 8.26 we can find a basis of V such that the matrix of N is upper-triangular with zeroes on the diagonal. Now the matrix of N^* is the conjugate transpose. Since $N^* = N$, the matrix of N must be zero, so the claim follows.

8. If $N^{\dim V - 1} \neq N^{\dim V}$, then $\dim \operatorname{null} N^{i-1} < \dim \operatorname{null} N^i$ for all $i \leq \dim V$ by proposition 8.5. By corollary $\dim \operatorname{null} N^{\dim V} = \dim V$, so the claim clearly follows.

9. $\operatorname{range} T^m = \operatorname{range} T^{m+1}$ implies $\dim \operatorname{null} T^m = \dim \operatorname{null} T^{m+1}$ i.e. $\operatorname{null} T^m = \operatorname{null} T^{m+1}$. From proposition 8.5 we get $\dim \operatorname{null} T^m = \dim \operatorname{null} T^{m+k}$ for all $k \geq 1$. Again this implies that $\dim \operatorname{range} T^m = \dim \operatorname{range} T^{m+k}$ for all $k \geq 1$, so the claim follows.

10. Let T be the operator defined in exercise 1. Clearly $\operatorname{null} T \cap \operatorname{range} T \neq \{0\}$, so the claim is false.

11. We have $\dim V = \dim \operatorname{null} T^n + \dim \operatorname{range} T^n$, so it's sufficient to prove that $\operatorname{null} T^n \cap \operatorname{range} T^n = \{0\}$. Let $v \in \operatorname{null} T^n \cap \operatorname{range} T^n$. Then we can find a $u \in V$ such that $T^n u = v$. From $0 = Tv = T^{n+1} u$ we see that $u \in \operatorname{null} T^{n+1} = \operatorname{null} T^n$ which implies that $v = T^n u = 0$.

12. From theorem 8.23 we have $V = \operatorname{null} T^{\dim V}$, so that $T^{\dim V} = 0$. Then let $T \in \mathcal{L}(\mathbb{R}^3)$ be the operator $T(a, b, c) = (-b, a, 0)$. Then clearly 0 is the only eigenvalue, but T is not nilpotent.

13. From $\operatorname{null} T^{n-2} \neq \operatorname{null} T^{n-1}$ and from proposition 8.5 we see that $\dim \operatorname{null} T^n \geq n - 1$. Assume that T has three different eigenvalues $0, \lambda_1, \lambda_2$. Then $\dim \operatorname{null} (T - \lambda_i I)^n \geq 1$, so from theorem 8.23

   $$n \geq \dim \operatorname{null} T^n + \dim \operatorname{null} (T - \lambda_1 I)^n + \dim \operatorname{null} (T - \lambda_2 I)^n \geq n - 1 + 1 + 1 = n + 1,$$

   which is impossible, so T has at most two different eigenvalues.

14. Let $T \in \mathcal{L}(\mathbb{C}^4)$ be defined by $T(a, b, c, d) = (7a, 7b, 8c, 8d)$ from the matrix of T it's easy to see that the characteristic polynomial is $(z - 7)^2 (z - 8)^2$.

15. Let d_1 be the multiplicity of the eigenvalue 5 and d_2 the multiplicity of the eigenvalue 6. Then $d_1, d_2 \geq 1$ and $d_1 + d_2 = n$. It follows that $d_1, d_2 \leq n - 1$, so that $(z - 5)^{d_1} (z - 6)^{d_2}$ divides $(z - 5)^{n-1} (z - 6)^{n-1}$. By the Cayley-Hamilton theorem $(T - 5I)^{d_1} (T - 6I)^{d_2} = 0$, so that $(T - 5I)^{n-1} (T - 6I)^{n-1} = 0$.

16. If every generalized eigenvector is an eigenvector, then by theorem 8.23 T has a basis of eigenvectors. If there's a generalized eigenvector that is not an eigenvector, then we have an eigenvalue λ such that $\dim \operatorname{null}(T - \lambda I)^{\dim V} \neq \dim \operatorname{null}(T - \lambda I)$. Thus if

$\lambda_1, \dots, \lambda_m$ are the eigenvalues of T, then $\sum_{i=1}^m \dim \operatorname{null}(T - \lambda_i I) < \dim V$, so there doesn't exists a basis of eigenvectors.

17. By lemma 8.26, choose a basis $(v_1, \dots, v_n)$ such that N has an upper-triangular matrix. Apply Gram-Schmidt orthogonalization to the basis. Then $Ne_1 = 0$ and assume that $Ne_i \in \operatorname{span}(e_1, \dots, e_i)$, so that

$$Ne_{i+1} = N\left(\frac{v_{i+1} - \langle v_{i+1}, e_1\rangle e_1 - \dots - \langle v_{i+1}, e_i\rangle e_i}{\|v_{i+1} - \langle v_{i+1}, e_1\rangle e_1 - \dots - \langle v_{i+1}, e_i\rangle e_i\|}\right) \in \operatorname{span}(e_1, \dots, e_{i+1}).$$

Thus N has an upper-triangular matrix in the basis $(e_1, \dots, e_n)$.

18. Continuing in the proof of lemma 8.30 up to $j = 4$ we see that $a_4 = -5/128$. So that

$$\sqrt{I+N} = I + \frac{1}{2}N - \frac{1}{8}N^2 + \frac{1}{16}N^3 - \frac{5}{128}N^4$$

19. Just replace the Taylor-polynomial in lemma 8.30 with the Taylor polynomial of $\sqrt[3]{1+x}$ and copy the proof of the lemma and theorem 8.32.

20. Let $p(x) = \sum_{i=0}^m a_i x^i$ be the minimal polynomial of T. If $a_0 \neq 0$, then $p(x) = x\sum_{i=1}^m a_i x^{i-1}$. Since T is invertible we must have $\sum_{i=1}^m a_i T^{i-1} = 0$ which contradicts the minimality of p. Hence $a_0 \neq 0$, so solving for I in $p(T)$ we get

$$-a_0^{-1}(a_1 + \dots + a_m T^{m-1})T = I.$$

Hence setting $q(x) = -a_0^{-1}(a_1 + \dots + a_m x^{m-1})$ we have $q(T) = T^{-1}$.

21. The operator defined by the matrix

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

is clearly an example.

22. Choose the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

It's easy to see that $A(A-I)^2 = 0$. Thus, the minimal polynomial divides $z(z-1)^2$. However, the whole polynomial is the only factor that annihilates A, so $z(z-1)^2$ is the minimal polynomial.

23. Let $\lambda_1, \ldots, \lambda_m$ be the eigenvalues of T with multiplicity $d_1, \ldots, d_m$. Assume that V has a basis of eigenvectors $(v_1, \ldots, v_n)$. Let $p(z) = \prod_{i=1}^m (z - \lambda_i)$. Then we get

$$p(T)v_i = \left(\prod_{j \neq i} (T - \lambda_j I) \right) (T - \lambda_i I)v_i = 0,$$

so that the minimal polynomial divides p and hence has no double root.

Now assume that the minimal polynomial has no double roots. Let the minimal polynomial be p and let $(v_1, \ldots, v_n)$ be a Jordan-basis of T. Let A be the largest Jordan-block of T. Now clearly $p(A) = 0$. However, by exercise 29, the minimal polynomial of $(A - \lambda I)$ is z^{m+1} where m is the length of the longest consecutive string of 1's that appear just above the diagonal, so the minimal polynomial of A is $(z - \lambda)^{m+1}$. Hence $m = 0$, so that A is a 1×1 matrix and the basis vector corresponding to A is an eigenvalue. Since A was the largest Jordan-block it follows that all basis vectors are eigenvalues (corresponding to a 1×1 matrix).

24. If V is a complex vector space, then V has a basis of eigenvectors. Now let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues. Now clearly $p = \prod_{i=1}^m (z - \lambda_i)$ annihilates T, so that the minimal polynomial being a factor of p doesn't have a double root.

Now assume that V is a real vector space. By theorem 7.25 we can find a basis of T such that T has a block diagonal matrix where each block is a 1×1 or 2×2 matrix. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues corresponding to the 1×1 blocks. Then $p(z) = \prod_{i=1}^m (z - \lambda_i)$ annihilates all but the 2×2 blocks of the matrix. Now it's sufficient to show that each 2×2 block is annihilated by a polynomial which doesn't have real roots.

By theorem 7.25 we can choose the 2×2 blocks to be of the form

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

where $b > 0$ and it's easy to see that

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} - 2a \begin{bmatrix} a & -b \\ b & a \end{bmatrix} + (a^2 + b^2) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0.$$

Clearly this polynomial has a negative discriminant, so the claim follows.

25. Let q be the minimal polynomial. Write $q = sp + r$, where $\deg r < \deg p$, so that $0 = q(T)v = s(T)p(T)v + r(T)v = r(T)v$. Assuming $r \neq 0$, we can multiply the both sides with the inverse of the highest coefficient of r yielding a monic polynomial r_2 of degree less than p such that $r_2(T)v = 0$ contradicting the minimality of p. Hence $r = 0$ and p divides q.

26. It's easy to see that no proper factor of $z(z-1)^2(z-3)$ annihilates the matrix

$$A = \begin{bmatrix} 3 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

so the it's minimal polynomial is $z(z-1)^2(z-3)$ which by definition is also the characteristic polynomial.

27. It's easy to see that no proper factor of $z(z-1)(z-3)$ annihilates the matrix

$$A = \begin{bmatrix} 3 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

but $A(A-I)(A-3I) = 0$, so the claim follows.

28. We see that $T(e_i) = e_{i+1}$ for $i < n$. Hence $(e_1, Te_1, \ldots, T^{n-1}e_1) = (e_1, \ldots, e_n)$ is linearly independent. Thus for any non-zero polynomial p of degree less than n we have $p(e_i) \neq 0$. Hence the minimal polynomial has degree n, so that the minimal polynomial equals the characteristic polynomial.

    Now from the matrix of T we see that $T^n(e_i) = -a_0 - a_1T(e_1) - \ldots - a_{n-1}T^{n-1}$. Set $p(z) = a_0 + a_1z + \ldots + a_{n-1}z^{n-1} + z^n$. Now $p(T)(e_1) = 0$, so by exercise 25 p divides the minimal polynomial. However, the minimal polynomial is monic and has degree n, so p is the minimal polynomial.

29. The biggest Jordan-block of N is of the form

$$\begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}.$$

    Now clearly $N^{m+1} = 0$, so the minimal polynomial divides z^{m+1}. Let $v_i, \ldots, v_{i+m}$ be the basis elements corresponding to the biggest Jordan-block. Then $N^m e_{i+m} = e_i \neq 0$, so the minimal polynomial is z^{m+1}.

30. Assume that V can't be decomposed into two proper subspaces. Then T has only one eigenvalue. If there is more than one Jordan-block, then Let $(v_1, \ldots, v_m)$ be the vectors corresponding to all but the last Jordan-block and $(v_{m+1}, \ldots, v_n)$ the

vectors corresponding to the last Jordan-block. Then clearly $V = \text{span}(v_1, \ldots, v_m) \oplus \text{span}(v_{m+1}, \ldots, v_n)$. Thus, we have only one Jordan-block and $T - \lambda I$ is nilpotent and has minimal polynomial $z^{\dim V}$ by the previous exercise. Hence T has minimal polynomial $(z - \lambda)^{\dim V}$.

Now assume that T has minimal polynomial $(z-\lambda)^{\dim V}$. If $V = U \oplus W$, let p_1 be the minimal polynomial of $T_{|U}$ and p_2 the minimal polynomial of $T_{|W}$. Now $(p_1 p_2)(T) = 0$, so that $(z - \lambda)^{\dim V}$ divides $p_1 p_2$. Hence $\deg p_1 p_2 = \deg p_1 + \deg p_2 \geq \dim V$, but $\deg p_1 + \deg p_2 \leq \dim U + \dim W = \dim V$. Thus $\deg p_1 p_2 = \dim V$. This means that $p_1(z) = (z - \lambda)^{\dim U}$ and $p_2(z) = (z - \lambda)^{\dim W}$. Now if $n = \max\{\dim U, \dim W\}$ we have $(T_{|U} - \lambda I)^n = (T_{|W} - \lambda I)^n = 0$. This means that $(T - \lambda I)^n = 0$ contradicting the fact that $(z - \lambda)^{\dim V}$ is the minimal polynomial.

31. Reversing the Jordan-basis simply reverses the order of the Jordan-blocks and each block needs to be replaced by its transpose.

9 Operators on Real Vector Spaces

1. Let a be the value in the upper-left corner. Then the matrix must have the form

$$\begin{bmatrix} a & 1-a \\ 1-a & a \end{bmatrix}$$

and it's easy to see that $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue 1.

2. The characteristic polynomial of the matrix is $p(x) = (x - a)(x - d) - bc$. Now the discriminant equals $(a + d)^2 - 4(ad - bc) = (a - d)^2 + 4bc$. Hence the characteristic polynomial has a root if and only if $(a - d)^2 + 4bc \geq 0$. If $p(x) = (x - \lambda_1)(x - \lambda_2)$, then $p(T) = 0$ implies that either $A - \lambda_1 I$ or $A - \lambda_2 I$ is not invertible hence A has an eigenvalue. If p doesn't have a root then assuming that $Av = \lambda v$ we get

$$0 = p(A)v = p(\lambda)v,$$

so that $v = 0$. Hence A has no eigenvalues. The claim follows.

3. See the proof of the next problem, though in this special case the proof is trivial.

4. First let λ be an eigenvalue of A. Assume that λ is not an eigenvalue of $A_1, \ldots, A_m$, so that $A_i - \lambda I$ is invertible for each i. Now let B be the matrix

$$\begin{bmatrix} (A_1 - \lambda I)^{-1} & & * \\ & \ddots & \\ 0 & & (A_m - \lambda I)^{-1} \end{bmatrix}.$$

It's now easy to verify that $B(A-\lambda I)$ is a diagonal matrix with all 1s on the diagonal, hence it's invertible. However, this is impossible, since $A - \lambda I$ is not. Thus, λ is an eigenvalue of some A_i.

Now let λ be an eigenvalue of A_i. Let T be the operator corresponding to $A - \lambda I$ in $\mathcal{L}(\mathbb{F}^n)$. Let A_i correspond to the columns $j, \ldots, j+k$ and let $(e_1, \ldots, e_n)$ be the standard basis. Let $(a_1, \ldots, a_k)$ be the eigenvector of A_i corresponding to λ. Set $v = a_1 e_j + \ldots + a_k e_{j+k}$. Then it's easy to see that T maps the space $\operatorname{span}(e_1, \ldots, e_{j-1}, v)$ to the space $\operatorname{span}(e_1, \ldots, e_{j-1})$. Hence T is not injective, so that λ is an eigenvalue of A.

5. The proof is identical to the argument used in the solution of exercise 2.

6. Let $(v_1, \ldots, v_n)$ be the basis with the respect to which T has the given matrix. Applying Gram-Schmidt to the basis the matrix trivially has the same form.

7. Choose a basis $(v_1, \ldots, v_n)$ such that T has a matrix of the form in theorem 9.10. Now if v_j corresponds to the second vector in a pair corresponding to a 2×2 matrix or corresponds to a 1×1 matrix, then $\operatorname{span}(v_1, \ldots, v_j)$ is an invariant subscape. The second posibility means that v_j corresponds to the first vector in a pair corresponding to a 2×2 matrix, so that $\operatorname{span}(v_1, \ldots, v_{j+1})$ is an invariant subspace.

8. Assuming that such an operator existed we would have basis $(v_1, \ldots, v_7)$ such that the matrix of T would have a matrix with eigenpair $(1, 1)$

$$\frac{\dim \operatorname{null}(T^2 + T + I)^{\dim V}}{2} = 7/2$$

times on the diagonal. This would contradict theorem 9.9 as $7/2$ is not a whole number. It follows that such an operator doesn't exist.

9. The equation $x^2 + x + 1 = 0$ has a solution in $\mathbb{C}$. Let λ be a solution. Then the operator corresponding to the matrix λI clearly is an example.

10. Let T be the operator in $\mathcal{L}(\mathbb{R}^{2k})$ corresponding to the block diagonal matrix where each block has characteristic polynomial $x^2 + \alpha x + \beta$. Then by theorem 9.9 we have

$$k = \frac{\dim \operatorname{null}(T^2 + \alpha T + \beta I)^{2k}}{2},$$

so that $\dim \operatorname{null}(T^2 + \alpha T + \beta I)^{2k}$ is even. However, the minimal polynomial of T divides $(T^2 + \alpha T + \beta I)^{2k}$ and has degree less than $2k$, so that $(T^2 + \alpha T + \beta I)^k = 0$. It follows that $\dim \operatorname{null}(T^2 + \alpha T + \beta I)^{2k} = \dim \operatorname{null}(T^2 + \alpha T + \beta I)^k$ and the claim follows.

11. We see that (α, β) is an eigenpair and from the nilpotency of $T^2 + \alpha T + \beta I$ and theorem 9.9 we have

$$\frac{\dim \operatorname{null}(T^2 + \alpha T + \beta I)^{\dim V}}{2} = (\dim V)/2$$

it follows that $\dim V$ is even. For the second part we know from the Cayley-Hamilton theorem that the minimal polynomial $p(x)$ of T has degree less than $\dim V$. Thus we have $p(x) = (x^2 + \alpha x + \beta)^k$ and from $\deg p \leq \dim V$ we get $k \leq (\dim V)/2$. Hence $(T + \alpha T + \beta I)^{(\dim V)/2} = 0$.

12. By theorem 9.9 we can choose a basis such that the matrix of T has the form of 9.10. Now we have the matrices [5] and [7] at least once on the diagonal. Assuming that T has an eigenpair we would have at least one 2×2 matrix on the diagonal. This is impossible as the diagonal has only length 3.

13. By proposition 8.5 $\dim \operatorname{null} T^{n-1} \geq n - 1$. Like in the previous exercise we know that the matrix [0] is at least $n - 1$ times on the diagonal and it's impossible to fit a 2×2 matrix in the only place left.

14. We see that A satisfies the polynomial $p(z) = (z - a)(z - d) - bc$ no matter if $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$. If $\mathbb{F} = \mathbb{C}$, then because p is monic, has degree 2 and annihilates A it follows that p is the characteristic polynomial.

    Assume then that $\mathbb{F} = \mathbb{R}$. Now if A has no eigenvalues, then p is the characteristic polynomial by definition. Assume then that $p(z) = (z - \lambda_1)(z - \lambda_2)$ which implies that $p(T) = (T - \lambda_1 I)(T - \lambda_2 I) = 0$. If neither $T - \lambda_1 I$ or $T - \lambda_2 I$ is injective, then by definition p is the minimal polynomial. Assume then that $T - \lambda_2 I$ is invertible. Then $\dim \operatorname{null}(T - \lambda_1 I) = 2$, so that $T - \lambda_1 I = 0$. Hence we get $c = b = 0$ and $a = d = \lambda_1$. Hence $p(z) = (z - \lambda_1)^2$, so that p is the characteristic polynomial by definition.

15. S is normal, so there's an orthonormal basis of V such that S has a block diagonal matrix with respect to the basis and each block is a 1×1 or 2×2 matrix and the 2×2 blocks have no eigenvalue (theorem 7.25). From $S^*S = I$ we see that each block A_i satisfy $A_i^* A_i = I$, so that they are isometries. Hence a 2×2 block is of the form

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

    We see that $T^2 + \alpha T + \beta I$ is not injective if and only if $A_i^2 + \alpha A_i + \beta I$ is not injective for some 2×2 matrix A_i. Hence $x^2 + \alpha x + \beta$ must equal the characteristic polynomial of some A_i, so by the previous exercise it's of the form

$$(x - \cos\theta)(z - \cos\theta) + \sin^2\theta = x^2 - 2x\cos\theta + \cos^2\theta + \sin^2\theta$$

    so that $\beta = \cos^2\theta + \sin^2\theta = 1$.

10 Trace and Determinant

1. The map $T \mapsto \mathcal{M}(T, (v_1, \ldots, v_n))$ is bijective and satisfies $ST \mapsto \mathcal{M}(ST, (v_1, \ldots, v_n)) = \mathcal{M}(S, (v_1, \ldots, v_n))\mathcal{M}(T, (v_1, \ldots, v_n))$. Hence

$$ST = I \Leftrightarrow \mathcal{M}(ST, (v_1, \ldots, v_n)) = \mathcal{M}(S, (v_1, \ldots, v_n))\mathcal{M}(T, (v_1, \ldots, v_n)) = I.$$

   The claim now follows trivially.

2. Both matrices represent an operator in $\mathcal{L}(\mathbb{F}^n)$. The claim now follows from exercise 3.23.

3. Choose a basis $(v_1, \ldots, v_n)$ and let $A = (a_{ij})$ be the matrix corresponding to the basis. Then $Tv_1 = a_{11}v_1 + \ldots + a_{n1}v_n$. Now $(v_1, 2v_2, \ldots, 2v_n)$ is also a basis and we have by our assumption $Tv = a_{11}v_1 + 2(a_{21}v_2 + \ldots + a_{n1}v_n)$. We thus get

$$a_{21}v_2 + \ldots + a_{n1}v_n = 2(a_{21}v_2 + \ldots + a_{n1}v_n),$$

   which implies $(a_{21}v_2 + \ldots + a_{n1}v_n = 0$. By linear independence we get $a_{21} = \ldots = a_{n1} = 0$, so that $Tv_1 = a_{11}v$. The claim clearly follows.

4. Follows directly from the definition of $\mathcal{M}(T, (u_1, \ldots, u_n), (v_1, \ldots, v_n))$.

5. Let $(e_1, \ldots, e_n)$ be the standard basis for $\mathbb{C}^n$. Then we can find an operator $T \in \mathcal{L}(\mathbb{C}^n)$ such that $\mathcal{M}(T, (e_1, \ldots, e_n)) = B$. Now T has an upper-triangular matrix corresponding to some basis $(v_1, \ldots, v_n)$ of V. Let $A = \mathcal{M}((v_1, \ldots, v_n), (e_1, \ldots, e_n))$. Then

$$A^{-1}BA = \mathcal{M}(T, (v_1, \ldots, v_n))$$

   which is upper-triangular. Clearly A is an invertible square matrix.

6. Let T be the operator corresponding to the matrix

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

   By theorem 10.11 we have $\text{trace}(T^2) = -2 < 0$.

7. Let $(v_1, \ldots, v_n)$ be a basis of eigenvectors of T and $\lambda_1, \ldots, \lambda_n$ be the corresponding eigenvalues. Then T has the matrix

$$\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}.$$

   Clearly $\text{trace}(T^2) = \lambda_1^2 + \ldots + \lambda_n^2 \geq 0$ by theorem 10.11.

8. Extend $v/\|v\|$ to an orthonormal basis $(v/\|v\|, e_1, \ldots, e_n)$ of V and let $A = \mathcal{M}(T, (v/\|v\|, e_1, \ldots, e_n))$. Let $a, a_1, \ldots, a_n$ denote the diagonal elements of A. Then we have

$$a_i = \langle Te_i, e_i\rangle = \langle \langle e_i, v\rangle w, e_i\rangle = \langle 0, e_i\rangle = 0,$$

so that

$$\text{trace}(T) = a = \langle Tv/\|v\|, v/\|v\|\rangle = \langle \langle v/\|v\|, v\rangle w, v/\|v\|\rangle = \langle w, v\rangle .$$

9. From exercise 5.21 we have $V = \text{null}\, P \oplus \text{range}\, P$. Now let $v \in \text{range}\, P$. Then $v = Pu$ for some $u \in V$. Hence $Pv = P^2w = Pw = v$, so that P is the identity on range P. Now choose a basis $(v_1, \ldots, v_m)$ for range P and extend it with a basis $(u_1, \ldots, u_n)$ of null P to get a basis for V. Then clearly the matrix for P in this basis consists of a diagonal matrix with 1s on part of the diagonal corresponding to the vectors $(v_1, \ldots, v_m)$ and 0s on the rest of the diagonal. Hence trace $P = \dim \text{range}\, P \geq 0$.

10. Let $(v_1, \ldots, v_n)$ be some basis of V and let $A = \mathcal{M}(T, (v_1, \ldots, v_n))$. Let $a_1, \ldots, a_n$ be the diagonal elements of A. By proposition 6.47 T^* has the matrix A^*, so that the diagonal elements of A^* are $\overline{a}_1, \ldots, \overline{a}_n$. By theorem 10.11 we have

$$\text{trace}(T^*) = \overline{a}_1 + \ldots + \overline{a}_n = \overline{a_1 + \ldots + a_n} = \overline{\text{trace}(T)}.$$

11. A positive operator is self-adjoint. By the spectral theorem we can find a basis $(v_1, \ldots, v_n)$ of eigenvectors of T. Then $A = \mathcal{M}(T, (v_1, \ldots, v_n))$ is a diagonal matrix and by positivity of T all the diagonal elements $a_1, \ldots, a_n$ are positive. Hence $a_1 + \ldots + a_n = 0$ implies $a_1 = \ldots = a_n = 0$, so that $A = 0$ and hence $T = 0$.

12. The trace of T is the sum of the eigenvalues. Hence $\lambda - 48 + 24 = 51 - 40 + 1$, so that $\lambda = 36$.

13. Choose a basis of $(v_1, \ldots, v_n)$ of V and let $A = \mathcal{M}(T, (v_1, \ldots, v_n))$. Then the matrix of cT is cA. Let $a_1, \ldots, a_n$ be the diagonal elements of A. Then we have

$$\text{trace}(cT) = ca_1 + \ldots + ca_n = c(a_1 + \ldots + a_n) = c\text{trace}(T).$$

14. The example in exercise 6 shows that this is false. For another example take $S = T = I$.

15. Choose a basis $(v_1, \ldots, v_n)$ for V and let $A = \mathcal{M}(T, (v_1, \ldots, v_n))$. It's sufficient to prove that $A = 0$. Now let $a_{i,j}$ be the element in row i, column j of A. Let B be the matrix with 1 in row j column i and 0 elsewhere. Then it's easy to see that in BA the only non-zero diagonal element is $a_{i,j}$. Thus $\text{trace}(BA) = a_{i,j}$. Let S be the operator corresponding to B. It follows that $a_{i,j} = \text{trace}(ST) = 0$, so that $A = 0$.

16. Let $T^*Te_i = a_1e_1 + \ldots + a_ne_n$. Then we have that $a_i = \langle T^*Te_i, e_i\rangle = \|Te_i\|^2$ by the ortogonality of $(e_1, \ldots, e_n)$. Clearly a_i is the ith diagonal element in the matrix $\mathcal{M}(T^*T, (e_1, \ldots, e_n))$, so that

$$\operatorname{trace}(T^*T) = \|Te_1\|^2 + \ldots + \|Te_n\|^2.$$

The second assertion follows immediately.

17. Choose a basis $(v_1, \ldots, v_n)$ such that T has an upper triangular matrix $B = (b_{ij})$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ on the diagonal. Now the ith element on the diagonal of B^*B is $\sum_{j=1}^n \bar{b}_{ji}b_{ji} = \sum_{j=1}^n |b_{ji}|^2$, where $|b_{ii}|^2 = |\lambda_i|^2$. Hence

$$\sum_{i=1}^n |\lambda_i|^2 \leq \sum_{i=1}^n \sum_{j=1}^n |b_{ji}|^2 = \operatorname{trace}(B^*B) = \operatorname{trace}(A^*A) = \sum_{k=1}^n \sum_{j=1}^n |a_{jk}|^2.$$

18. Positivity and definiteness follows from exercise 16 and additivity in the first slot from corollary 10.12. Homogeneity in the first slot is exercise 13 and by exercise 10

$$\langle S, T\rangle = \operatorname{trace}(ST^*) = \overline{\operatorname{trace}((ST^*)^*)} = \overline{\operatorname{trace}(TS^*)} = \overline{\langle T, S\rangle}.$$

It follows that the formula defines an inner-product.

19. We have $0 \leq \langle Tv, Tv\rangle - \langle T^*v, T^*v\rangle = \langle (T^*T - TT^*)v, v\rangle$. Hence $T^*T - TT^*$ is a positive operator (it's clearly self-adjoint), but $\operatorname{trace}(T^*T - TT^*) = 0$, so all the eigenvalues are 0. Hence $T^*T - TT^* = 0$ by the spectral theorem, so T is normal.

20. Let $\dim V = n$ and write $A = \mathcal{M}(T)$. Then we have

$$\det cT = \det cA = \sum_\sigma ca_{\sigma(1),1} \cdots ca_{\sigma(n),n} = c^n \det A = c^n \det T$$

21. Let $S = I$ and $T = -I$. Then we have $0 = \det(S + T) \neq \det S + \det T = 1 + 1 = 2$.

22. We can use the result of the next exercise which means that we only need to prove this for the complex case. Let $T \in \mathcal{L}(\mathbb{C}^n)$ be the operator corresponding to the matrix A. Now each block A_j on the diagonal corresponds to some basis vectors $(e_i, \ldots, e_{i+k})$. These can be replaced with another set of vectors, $(v_i, \ldots, v_{i+k})$, spanning the same subspace such that A_j in this new basis is upper-triangular. We have $T(\operatorname{span}(v_i, \ldots, v_{i+k})) \subset \operatorname{span}(v_i, \ldots, v_{i+k})$, so after doing this for all blocks we can assume that A is upper-triangular. The claim follows immediately.

23. This follows trivially from the formula of trace and determinant for a matrix, because the formulas only depends on the elements of the matrix.

24. Let $\dim V = n$ and let $A = \mathcal{M}(T)$, so that $B = \mathcal{M}(T)$ equals the conjugate transpose of A i.e. $b_{ij} = \overline{a_{ji}}$. Then we have

$$\begin{aligned}\overline{\det T} &= \overline{\det A} = \overline{\sum_{\sigma} a_{\sigma(1),1} \cdots a_{\sigma(n),n}} \\ &= \overline{\sum_{\sigma} a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}} = \sum_{\sigma} \overline{a}_{1,\sigma(1)} \cdots \overline{a}_{n,\sigma(n)} \\ &= \sum_{\sigma} b_{\sigma(1),1} \cdots b_{\sigma(n),n} = \det B = \det T^*.\end{aligned}$$

The first equality on the second line follows easily that every term in the upper sum is represented by a term in the lower sum and vice versa. The second claim follows immediately from theorem 10.31.

25. Let $\Omega = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 < 1\}$ and let T be the operator $T(x, y, z) = (ax, by, cz)$. It's easy to see that $T(\Omega)$ is the ellipsoid. Now Ω is a circle with radius 1. Hence

$$|\det T|\text{volume}(\Omega) = abc \times \frac{4}{3}\pi = \frac{4}{3}\pi abc,$$

which is the volume of an ellipsoid.