Sheldon Axler's Linear Algebra Done Right - A partial solutions manual

George Lydakis

June 1, 2024

Contents

Chapter 5

Eigenvalues, Eigenvectors and Invariant Subspaces

5.A Invariant Subspaces

Exercise 24
Suppose A is an n-by-n matrix with entries in $\mathbf{F}$. Define $T \in L(\mathbf{F^n})$ by $Tx = Ax$, where elements of $\mathbf{F^n}$ are thought of as n-by-1 column vectors.
(a) Suppose the sum of the entries in each row of A equals 1. Prove that 1 is an eigenvalue of T.
(b) Suppose the sum of the entries in each column of A equals 1. Prove that 1 is an eigenvalue of T.

Solution.
(a) We need to show that there exists non-zero $x \in \mathbf{F^n}$ such that $Tx = x$. We know that $Tx = Ax$. Let:

$$x = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Then:

$$Ax = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n A_{1i} \\ \sum_{i=1}^n A_{2i} \\ \vdots \\ \sum_{i=1}^n A_{ni} \end{pmatrix}$$

However, we know that the sum of each row of A is 1. This means that $\forall j, 1 \leq j \leq n, \sum_{i=1}^n A_{ji} = 1$. It is then clear that:

$$Ax = \begin{pmatrix} \sum_{i=1}^n A_{1i} \\ \sum_{i=1}^n A_{2i} \\ \vdots \\ \sum_{i=1}^n A_{ni} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = x$$

, which means that $Tx = Ax = x$, therefore 1 is an eigenvalue of T with x being a corresponding eigenvector.
(b) Again, we need to show that there exists non-zero $x \in \mathbf{F^n}$ such that $Tx = x$. If $T^{'}$ is the dual map of T, we know that $M(T') = A^T$ (with respect to the dual basis of the standard basis of $\mathbf{F^n}$. Since every column of A has a sum of 1, by the definition of the transpose we know that every row of A^T has a sum of 1. Therefore, by part (a) we know that the vector ϕ such that $M(\phi) = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$ with respect to the dual

basis of the standard basis of $\mathbf{F^n}$ is an eigenvector of T' with 1 being the corresponding eigenvalue. This in turn means that:

$$T^{'}\phi = \phi \implies T^{'}\phi - \phi = 0 \implies \phi(T(v)) - \phi(v) = 0, \forall v \in \mathbf{F^n} \implies \phi(T(v) - v) = 0, \forall v \in \mathbf{F^n} \implies$$
$$\phi((T - I)(v)) = 0, \forall v \in \mathbf{F^n}$$

We observe that this means that $\phi \in (range(T - I))^0$, that is, ϕ is in the annihilator of $range(T - I)$. Since ϕ is not the zero functional, we conclude that $dim((range(T - I))^0) \geq 1$. Then:

$$dim(range(T-I)) + dim((range(T-I))^0) = dim\mathbf{F^n} \implies dim(range(T-I)) = dim\mathbf{F^n} - dim((range(T-I))^0) \implies$$
$$dim(range(T - I)) < dim\mathbf{F^n}$$

, where the last inequality holds since we are subtracting at least 1 from $dim\mathbf{F^n}$. Therefore, $T - I$ is not surjective, which means that 1 is an eigenvalue of T.

5.B Eigenvectors and Upper-Triangular Matrices

Exercise 3
Suppose $T \in L(V)$ and $T^2 = I$ and -1 is not an eigenvalue of T. Prove that $T = I$.

Solution.
Observe that by the definition of linear map composition and the distributive property, we have:

$$(T - I)(T + I)(v) = T^2(v) + TI(v) - IT(v) - I^2(v) = T^2(v) + T(v) - T(v) - v = T^2(v) - v$$

, for all $v \in V$. Furthermore, we know that $T^2 = I$, therefore:

$$(T - I)(T + I)(v) = v - v = 0$$

Since -1 is not an eigenvalue of T, $T(v) \neq -v$ for all (nonzero) $v \in V$, which means that $T(v) + v = (T + I)(v) \neq 0$ for all $v \in V$. If $V = \{0\}$ the exercise is trivial since $T(0) = 0$ for any linear map. In the non-trivial case, there exists at least one nonzero $v \in V$, for which $(T + I)(v) = w \neq 0$. But:

$$(T - I)(T + I)(v) = (T - I)(w) = 0$$

, which means that $T(w) = w$ for a non-zero w, therefore 1 is an eigenvalue of T.

Exercise 9
Suppose V is finite-dimensional, $T \in L(V)$ and $v \in V$ with $v \neq 0$. Let p be a nonzero polynomial of smallest degree such that $p(T)(v) = 0$. Prove that every zero of p is an eigenvalue of T.

Solution.
Let λ be a zero of p, i.e. $p(\lambda) = 0$. p can be written as:

$$p(z) = (z - \lambda)q(z)$$

This means that $p(T)(u) = (T - \lambda I)q(T)(u)$ for all $u \in V$. q is of smaller degree than p, which means that $q(T)(v) \neq 0$, since by definition p has the smallest degree such that $p(T)(v) = 0$. Then let $w = q(T)(v) \neq 0$ which yields that:

$$p(T)(v) = (T - \lambda I)(w) = 0$$

, and since w is nonzero we have that λ is an eigenvalue of T. Since λ was selected as any zero of p, we have that all zeros of p are eigenvalues of T.

Exercise 13
Suppose W is a complex vector space and $T \in L(W)$ has no eigenvalues. Prove that every subspace of W invariant under T is either 0 or infinite-dimensional.

Solution.
If W is $\{0\}$ the exercise is trivial, therefore from now on we assume that there exists $v \in W, v \neq 0$. Obviously, $\{0\}$ is invariant under T. Let $U \neq \{0\}$ be a subspace of W invariant under T. We need to show that U is infinite-dimensional. We'll do this by proving that for any given integer m, U contains a linearly independent list of length m (by a previous result in the book this is equivalent to U being infinite-dimensional, the intuition is that no finite spanning list can exist since it should be at least as long as any linearly independent list).
For any positive integer m, consider the list:

$$v, Tv, T^2(v), \dots, T^m(v)$$

Since U is invariant under T, all elements of this list are also in U. Assume that the list is linearly dependent. Then, we can follow the structure of Axler's proof of the result "Every operator on a finite-dimensional non-zero complex vector space has an eigenvalue" to prove that T must have an eigenvalue, which is a contradiction. Therefore the list is linearly independent, and since this is true for any m, U must be infinite-dimensional.

5.C Eigenspaces and Diagonal Matrices

Exercise 7
Suppose $T \in L(V)$ has a diagonal matrix A with respect to some basis of V and that $\lambda \in \mathbf{F}$. Prove that λ appears on the diagonal of A precisely $\dim E(\lambda, T)$ times.

Solution.
Since T is diagonalizable, there exists a basis of V, $v_1, v_2, \dots, v_n$ such that all v_i are eigenvectors of T. Furthermore, by the definition of the matrix of T with respect to that basis, it's true that:

$$Tv_i = A_{ii} v_i$$

,and A_{ii} are precisely the eigenvalues of T (i.e. all of them are eigenvalues and there are no other eigenvalues). Now, let $\lambda \in \mathbf{F}$. If λ is not an eigenvalue of T, then $\dim E(\lambda, T) = 0$, and from the previous observation, λ cannot equal any of the A_{ii}, which means that it appears zero times on the diagonal of $M(T)$, thus the claim is true.
If, on the other hand, λ is an eigenvalue of T, then it must appear on the diagonal. Then, let $I = \{i \in \{1, 2, \dots, n\}$ such that $A_{ii} = \lambda\}$, i.e. the set of all indices for which the corresponding element of the diagonal equals λ. Let also l be the list of v_k such that $k \in I$. We observe that all of the v_k are in $E(\lambda, T)$, since they are eigenvectors corresponding to λ. Furthermore, they are linearly independent (sub-list of basis). If we assume that they not a spanning list of $E(\lambda, T)$, then there must exist a $v \in E(\lambda, T)$ linearly independent from all of them. But this v cannot belong in the basis of V, because then it would have been included in l. This means that the basis is not a spanning list of V, contradiction. Therefore, l spans $E(\lambda, T)$, and, being linearly independent, is a basis of it. Therefore its length, equal to the appearances of λ in the diagonal, is equal to $\dim E(\lambda, T)$.

Exercise 13
Find $R, T \in L(\mathbf{F^4})$ such that R, T each have 2, 6, 7 as eigenvalues, R, T have no other eigenvalues, and there does not exist an invertible operator $S \in L(\mathbf{F^4})$ such that $R = S^{-1}TS$.

Solution.
The example itself is not so much of interest here as is the reasoning for finding it. Our starting point will be to explore if this would be possible for R diagonalizable and T not diagonalizable. Since R is

diagonalizable, there exists a basis $v_1, v_2, v_3, v_4 \in \mathbf{F^4}$ all of which are eigenvectors of R. S would then be fully defined by its values $Sv_1 = w_1, Sv_2 = w_2, Sv_3 = w_3, Sv_4 = w_4$. Since S should be invertible, this list should be a basis of V too. It would also be true that:

$$R = S^{-1}TS \implies T = SRS^{-1}$$

Then:

$$Tw_i = SRS^{-1}(w_i) = SR(v_i) = S(\lambda_i v_i) = \lambda_i w_i$$

, where we used the fact that $S^{-1}w_i = v_i$ and assumed that λ_i is the eigenvalue of R corresponding to v_i. But then this means that each w_i is an eigenvector of T, therefore that there exists a basis of V consisting of eigenvectors of T, therefore T is diagonalizable, contradiction.
It suffices therefore to find R diagonalizable and T not diagonalizable. One such example is

$$R = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix}, T = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix}$$

, where it's clear that R is diagonalizable (it's diagonal wrt. the standard basis of $\mathbf{F^4}$) and that both R, T have exactly 2, 6, 7 as eigenvalues. It remains to be proved that T is indeed not diagonalizable. To see this, let's solve for the eigenvectors of T:

$$Tv = \lambda v \implies \begin{pmatrix} 2 & 1 & 1 & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

$$\implies \begin{pmatrix} 2x_1 + x_2 + x_3 + x_4 \\ 2x_2 \\ 6x_3 \\ 7x_4 \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \lambda x_3 \\ \lambda x_4 \end{pmatrix}$$

, for $\lambda = 7$ we get that $x_4 \in \mathbf{F}$, $x_3 = 0$, $x_2 = 0$ and $x_1 = x_4/5$. $\begin{pmatrix} 1/5 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ is then a basis of $E(7, T)$. For $\lambda = 6$ we get that $x_3 \in \mathbf{F^4}, x_4 = x_2 = 0$ and $x_1 = x_3/4$. $\begin{pmatrix} 1/4 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ is then a basis of $E(6, T)$. For $\lambda = 2$, we get that $x_2 = x_3 = x_4 = 0, x_1 \in \mathbf{F}$. $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ is then a basis of $E(2, T)$. We can then see that there exist only three linearly independent vectors of T, therefore not enough for a basis of $\mathbf{F^4}$, therefore T cannot be diagonalizable.

Chapter 6

Inner product spaces

6.A Inner Products and Norms

Exercise 3
Suppose $\mathbf{F} = \mathbf{R}$ and $V \neq \{0\}$. Replace the positivity condition (which states that $\langle v, v\rangle \geq 0$ for all $v \in V$) in the definition of an inner product (6.3) with the condition that $\langle v, v\rangle > 0$ for some $v \in V$. Show that this change in the definition does not change the set of functions from $V \times V$ to $\mathbf{R}$ that are inner products on V.

Solution.
Let S_1 be the set of functions from $V \times V$ to $\mathbf{R}$ that are inner products on V that are inner products according to 6.3. Let S_2 be the set of functions from $V \times V$ to $\mathbf{R}$ that are inner products on V according to the definition of this exercise. Both of the definitions include the properties that $\langle v, v\rangle = 0$ iff $v = 0$, and that the inner product is linear wrt. the first slot. Furthermore, since $\mathbf{F} = \mathbf{R}$, it's true that $\langle v, u\rangle = \langle u, v\rangle$ for any $u, v \in V$. These are facts we'll use throughout the proof.
We need to show that $S_1 = S_2$.
First, let f be an inner product belonging in S_1. Then, by the definition of S_1 it's true that $\langle u, u\rangle \geq 0$ for all $u \in V$ (here $\langle , \rangle$ is the inner product according to f). Since our $V \neq \{0\}$, there exists at least one $v \in V$ that is not zero, and by the properties I listed previously, it must be true that $\langle v, v\rangle \neq 0$, which then means that $\langle v, v\rangle > 0$. We observe then that the condition for f to belong in S_2 holds (along with the other two properties S_1, S_2 have in common), therefore $f \in S_2$, therefore $S_1 \subset S_2$.
Now, let f be an inner product belonging in S_2. Then, there exists $v \in V$ such that $v \neq 0, \langle v, v\rangle > 0$ ($\langle , \rangle$ now refers to the inner product according to f). We need to show that for any $u \in V, \langle u, u\rangle \geq 0$. Let $u \in V, u \neq v$ be a vector of V (it clearly holds for $u = v$).

- If $u = 0$, then clearly $\langle u, u\rangle = 0 \geq 0$.
- If $u \neq 0$ and u, v are linearly dependent, there exist $a, b \in \mathbf{F}$ not both 0 such that $au + bv = 0$. Furthermore, a cannot be 0 since that would imply $v = 0$. Then $u = -\frac{b}{a}v = cv$, which yields $\langle u, u\rangle = \langle cv, cv\rangle = c\langle v, cv\rangle = c\bar{c}\langle v, v\rangle = c^2\langle v, v\rangle$ (since $\mathbf{F} = \mathbf{R}$), so $\langle u, u\rangle \geq 0$ as a product of non-negative numbers.
- If $u \neq 0$ and u, v linearly independent, then $au + bv \neq 0$ whenever at least one of a, b is not 0, which is equivalent to $\langle au + bv, au + bv\rangle \neq 0$. We have that:

$$\langle au + bv, au + bv\rangle = a^2\langle u, u\rangle + ab\langle u, v\rangle + ab\langle v, u\rangle + b^2\langle v, v\rangle = a^2\langle u, u\rangle + 2ab\langle u, v\rangle + b^2\langle v, v\rangle$$

  Suppose now that $\lambda = \langle u, u\rangle, \lambda < 0$ and set $b = 1$. Then the expression above equals:

$$a^2\lambda + 2a\langle u, v\rangle + \langle v, v\rangle$$

If we consider this to be a second-degree polynomial with respect to a, the discriminant is:

$\Delta = 4(\langle u, v\rangle)^2 - 4\lambda\langle v, v\rangle$

Since $\lambda < 0, \langle v, v\rangle > 0$ we have that $-4\lambda\langle v, v\rangle > 0$, and $4(\langle u, v\rangle)^2 \geq 0$, therefore the discriminant is positive, and thus there exists a such that this quantity is 0, which means that $\langle au + v, au + v\rangle$ is 0, which is a contradiction. Therefore, $\langle u, u\rangle \geq 0$.

We've shown then that f satisfies the property of $\langle v, v\rangle \geq 0$ for all $v \in V$, and therefore $f \in S_1$, which means $S_2 \subset S_1$, and since $S_1 \subset S_2$, $S_1 = S_2$, which concludes the proof.

Exercise 4
Suppose V is a real inner product space. (a) Show that $\langle u+v, u-v\rangle = ||u||^2 - ||v||^2$ for every $u, v \in V$. (b) Show that if $u, v \in V$ have the same norm, then $u + v$ is orthogonal to $u - v$. (c) Use part (b) to show that the diagonals of a rhombus are perpendicular to each other.

Solution.
(a) We have that: $\langle u + v, u - v\rangle = \langle u, u\rangle - \langle u, v\rangle + \langle v, u\rangle - \langle v, v\rangle = \langle u, u\rangle - \langle v, v\rangle = ||u||^2 - ||v||^2$. For the derivation we've used the linearity of inner products, the definition of norms and the fact that we're in a real vector space, therefore we can dispense with conjugates.
(b) This is trivially derived from part (a), since $||u||^2 = ||v||^2$ implies that $\langle u + v, u - v\rangle = 0$, which means that $u + v, u - v$ are orthogonal.
(c) A rhombus is defined as a quadrilateral whose sides are all of equal length. If we assume that two adjacent sides of a rhombus are described by the vectors u, v, then the diagonals of the rhombus are $u - v, u + v$. By part (b) and the definition of the rhombus, $||u||^2 = ||v||^2$, therefore $u + v, u - v$ are orthogonal, that is, perpendicular to each other.

Exercise 6
Suppose $u, v \in V$. Prove that $\langle u, v\rangle = 0$ iff

$$||u|| \leq ||u + \alpha v||$$

for all $\alpha \in \mathbf{F}$.

Solution.
$\Longrightarrow$: If $\langle u, v\rangle = 0$, then for any α $u, \alpha v$ are orthogonal and we can apply the Pythagorean theorem:

$$||u + \alpha v||^2 = ||u||^2 + ||\alpha v||^2 \geq ||u||^2 \implies ||u + \alpha v|| \geq ||u||$$

$\Longleftarrow$: If $||u|| \leq ||u + \alpha v||$ for all $\alpha \in \mathbf{F}$, then we have the following:

$$||u||^2 \leq ||u + \alpha v||^2 \implies \langle u, u\rangle \leq \langle u + \alpha v, u + \alpha v\rangle \implies \langle u, u\rangle \leq \langle u, u\rangle + \langle u, \alpha v\rangle + \langle \alpha v, u\rangle, \langle \alpha v, \alpha v\rangle$$

$$\implies 0 \leq \overline{\alpha}\langle u, v\rangle + \alpha\langle v, u\rangle + \alpha\overline{\alpha}\langle v, v\rangle$$

If $v = 0$, then $\langle u, v\rangle = 0$ for any v. Otherwise, set $\alpha = -\frac{\langle u,v\rangle}{||v||^2}$ to obtain:

$$0 \leq -\frac{\overline{\langle u, v\rangle}\langle u, v\rangle}{||v||^2} - \frac{\langle u, v\rangle\langle v, u\rangle}{||v||^2} + \frac{\langle u, v\rangle\overline{\langle u, v\rangle}}{||v||^2} \implies 0 \leq -|\langle u, v\rangle|^2 - |\langle u, v\rangle|^2 + |\langle u, v\rangle|^2 \implies 0 \leq -|\langle u, v\rangle|^2$$

This implies of course that $\langle u, v\rangle = 0$, qed.

Exercise 12
Prove that $(x_1+\ldots+x_n)^2 \leq n(x_1^2+\ldots+x_n^2)$ for all positive integers n and all real numbers $x_1,\ldots,x_n$.

Solution.
Let $u = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbf{R}^n$ and $v = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbf{R}^n$. Observe then, that:

$$\langle u,v\rangle = x_1+x_2+\ldots+x_n, ||u|| = \sqrt{x_1^2+\ldots+x_n^2}, ||v|| = \sqrt{1^2+\ldots+1^2} = \sqrt{n}$$

By the Cauchy-Schwarz inequality, we have that:

$$|\langle u,v\rangle| \leq ||u||||v|| \implies |x_1+x_2+\ldots+x_n| \leq \sqrt{x_1^2+\ldots+x_n^2}\sqrt{n} \implies (x_1+\ldots+x_n)^2 \leq n(x_1^2+\ldots+x_n^2)$$

Exercise 13
Suppose u,v are nonzero vectors in $\mathbf{R}^2$. Prove that

$$\langle u,v\rangle = ||u||||v||cos(\theta)$$

, where θ is the angle between u,v (thinking of u,v as arrows with initial point at the origin. *Hint:* Draw the triangle formed by $u,v,u-v$ and then use the law of cosines.

Solution.
By the law of cosines applied on the indicated triangle (which may be degenerate if u,v lie on the same line, but is well defined since they are nonzero), and specifically on angle θ whose opposite side is $u-v$, we have that:

$$||u-v||^2 = ||u||^2+||v||^2-2||u||||v||cos(\theta) \implies \langle u-v,u-v\rangle = \langle u,u\rangle+\langle v,v\rangle-2||u||||v||cos(\theta)$$
$$\implies \langle u,u\rangle-\langle u,v\rangle-\langle v,u\rangle+\langle v,v\rangle = \langle u,u\rangle+\langle v,v\rangle-2||u||||v||cos(\theta) \implies -2\langle u,v\rangle = -2||u||||v||cos(\theta)$$
$$\implies \langle u,v\rangle = ||u||||v||cos(\theta)$$

Exercise 14
The angle between two vectors (thought of as arrows with initial point at the origin) in $\mathbf{R}^2$ or $\mathbf{R}^3$ can be defined geometrically. However, geometry is not as clear in $\mathbf{R}^n$ for $n>3$. Thus the angle between two nonzero vectors $x,y\in\mathbf{R}^n$ is defined to be:

$$arccos\frac{\langle x,y\rangle}{||x||||y||}$$

, where the motivation for this definition comes from the previous exercise. Explain why the Cauchy-Schwarz Inequality is needed to show that this definition makes sense.

Solution.
As we now, the cosine of an angle is a number in the interval $[-1,1]$. Therefore, the function *arccos* has $[-1,1]$ as its domain. This means that the quantity $\frac{\langle x,y\rangle}{||x||||y||}$ must always be between $[-1,1]$ if we want its *arccos* to be defined. By the Cauchy-Schwarz inequality we have that:

$$|\langle x,y\rangle| \leq ||x||||y|| \implies -||x||||y|| \leq \langle x,y\rangle \leq ||x||||y|| \implies -1 \leq \frac{\langle x,y\rangle}{||x||||y||} \leq 1$$

, where we divide by the magnitude of x,y since these are nonzero vectors. Therefore, the argument of *arccos* is always in its domain and the definition makes sense.

It's also worth noting that from the "original" geometrical perspective, the cosine of the angle of two vectors is 1 when one of the vectors is a positive scalar multiple of the other and -1 when one of the vectors is a negative multiple of the other. Again by the Cauchy-Schwarz inequality, we know that the inequality becomes an equality precisely when one of the vectors is a scalar multiple of the other, which makes the corresponding argument to *arccoss* -1 or 1. Furthermore, if said scalar is positive, $\frac{\langle x,y\rangle}{||x||||y||} = \frac{\langle x,\alpha x\rangle}{||x||||\alpha x||} = \frac{\alpha||x||^2}{\alpha||x||||x||} = 1$, and if said scalar is negative $||\alpha x|| = -\alpha||x||$, so the same derivation yields that this quantity is -1. Therefore the angle's cosine is -1 or 1 exactly when we would expect it to be, and it's also 0 exactly when x, y are orthogonal.

Exercise 17
Prove or disprove: there is an inner product on $\mathbf{R}^2$ such that the associated norm is given by

$$||(x,y)|| = max\{x,y\}$$

for all $(x,y) \in \mathbf{R}^2$

Solution.
If such an inner product existed then consider the vector $v = (-1,0)$. For this vector, $||v|| = max\{-1,0\} = 0$. Then, $\langle v,v\rangle = ||v||^2 = 0$, but the vector is nonzero. This contradicts the definition of an inner product, which states that $\langle v,v\rangle = 0$ iff $v = 0$. Therefore there is no such inner product on $\mathbf{R}^2$.

Exercise 18
Suppose $p > 0$. Prove that there is an inner product on $\mathbf{R}^2$ such that the associated norm is given by

$$||(x,y)|| = (x^p + y^p)^{1/p}$$

for all $(x,y) \in \mathbf{R}^2$ if and only if $p = 2$

Solution.
$\Longleftarrow$: Suppose $p = 2$. Then $||(x,y)|| = (x^p + y^p)^{1/p} = \sqrt{x^2+y^2}$. As we know, this is the Euclidean norm on $\mathbf{R}^2$, which corresponds to the inner product $\langle (x_1,y_1),(y_1,y_2)\rangle = x_1y_1 + x_2y_2$. Therefore, an inner product associated with this norm does exist on $\mathbf{R}^2$.
$\Longrightarrow$: Suppose now that $p \neq 2, p > 0$.
If p is odd, i.e. $p = 2k+1$ for $k \geq 0$, then consider the vector $u = (-1,1)$. Its norm would be $||(-1,1)|| = ((-1)^{2k+1} + (1)^{2k+1})^{1/(2k+1)} = (-1+1)^{1/(2k+1)} = 0$. However, $(-1,-1) \neq (0,0)$, and therefore if there was an inner product associated with this norm it would be the case that $\langle u,u\rangle = ||u||^2 = 0$ with $u \neq 0$, which contradicts the definitiness property of inner products.
If p is even, i.e. $p = 2k$ for $k > 1$, we have the following. Firstly, problem 6 (for which I "owe" the proof) states that $\langle u,v\rangle = 0 \iff ||u|| \leq ||u+\alpha v||$ for all $\alpha \in \mathbf{F}$. Let us consider what would happen if there was an inner product associated with this norm. Let $u = (0,1), v = (1,0)$. Then, for any $\alpha \in \mathbf{F}$ we have that:

$$||u|| \leq ||u+\alpha v|| \iff (0^{2k} + 1^{2k})^{1/2k} \leq (a^{2k} + 1^{2k})^{1/2k} \iff 1 \leq a^{2k} + 1$$

, where we can square both sides and have an equivalence since they're both positive ($2k$ is even, thus $a^{2k} \geq 0$). Furthermore, this last inequality holds for any α since $2k$ is even. Therefore, by problem 6, $\langle (0,1),(1,0)\rangle = 0$. Since we have a norm and an associated inner product, the Pythagorean theorem must hold:

$$||(0,1)+(1,0)||^2 = ||(0,1)||^2 + ||(1,0)||^2 \implies ||(1,1)||^2 = ||(0,1)||^2 + ||(1,0)||^2$$
$$\implies (1^{2k} + 1^{2k})^{2/2k} = (0^{2k} + 1^{2k})^{2/2k} + (1^{2k} + 0^{2k})^{2/2k} \implies (2)^{1/k} = 1 + 1 = 2$$

This implies, however, that $k = 1$, which would yield $p = 2$, which is a contradiction. This completes our proof of the " $\Longrightarrow$ " direction, and thus also of the equivalence we needed to prove.

Exercise 19
Suppose V is a real inner product space. Prove that:

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

for all $u, v \in V$.

Solution.
We have that (setting N to be the numerator of the exercise's fraction):

$$N = ||u+v||^2 - ||u-v||^2 = ||u+v||^2 + ||u-v||^2 - 2||u-v||^2$$

By the parallelogram equality, we continue with:

$$N = 2(||u||^2 + ||v||^2) - 2\langle u-v, u-v\rangle = 2\langle u,u\rangle + 2\langle v,v\rangle - 2\langle u,u\rangle + 2\langle u,v\rangle + 2\langle v,u\rangle - 2\langle v,v\rangle = 4\langle u,v\rangle$$

, where the last equality comes from the fact that in a real vector space $\langle u,v\rangle = \langle v,u\rangle$. Our numerator is thus $N = 4\langle u,v\rangle$, and therefore $\frac{N}{4} = \langle u,v\rangle$, which is what we are asked to prove.

Exercise 20
Suppose V is a complex inner product space. Prove that

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

for all $u, v \in V$.

Solution.
Let us set $A = ||u+v||^2 - ||u-v||^2$ and $B = ||u+iv||^2 i - ||u-iv||^2 i$, in which case we need to prove that $\langle u,v\rangle = \frac{A+B}{4}$. We have that:

$$A = \langle u+v, u+v\rangle - \langle u-v, u-v\rangle = \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle - (\langle u,u\rangle - \langle u,v\rangle - \langle v,u\rangle + \langle v,v\rangle)$$
$$= 2\langle u,v\rangle + 2\langle v,u\rangle$$

$$B = i(\langle u+iv, u+iv\rangle - \langle u-iv, u-iv\rangle) = i(\langle u,u\rangle + \langle u,iv\rangle + \langle iv,u\rangle + \langle iv,iv\rangle - (\langle u,u\rangle - \langle u,iv\rangle - \langle iv,u\rangle + \langle iv,iv\rangle))$$
$$= i(2\langle u,iv\rangle + 2\langle iv,u\rangle) = i(2\bar{i}\langle u,v\rangle + 2i\langle v,u\rangle) = i(-2i\langle u,v\rangle + 2i\langle v,u\rangle) = 2\langle u,v\rangle - 2\langle v,u\rangle$$

Then $\frac{A+B}{4} = \frac{2\langle u,v\rangle + 2\langle v,u\rangle + 2\langle u,v\rangle - 2\langle v,u\rangle}{4} = \frac{4\langle u,v\rangle}{4} = \langle u,v\rangle$, qed.

Exercise 21
A norm on a vector space U is a function $|| \ || : U \to [0, \infty)$ such that $||u|| = 0$ if and only if $u = 0, ||\alpha u|| = |\alpha|||u||$ for all $\alpha \in \mathbf{F}$ and all $u \in U$, and $||u+v|| \leq ||u|| + ||v||$ for all $u, v \in U$. Prove that a norm satisfying the parallelogram equality comes from an inner product (in other words, show that if $|| \ ||$ is a norm on U satisfying the parallelogram equality, then there is an inner product $\langle , \rangle$ on U such that $||u|| = \langle u, u\rangle^{1/2}$ for all $u \in U$).

Solution.
We will begin by assuming that U is a real vector space. Define:

$$\langle x, y\rangle = \frac{1}{4}(||x+y||^2 - ||x-y||^2)$$

We will now prove that this function of x, y defines an inner product on U for which it is true that $||x||^2 = \langle x, x\rangle$ for all $x \in U$. From now on, we'll refer to the three properties of the norm assumed in the exercise, and in the same order as presented, as property 1, 2, 3 respectively. Firstly:

$$\langle x, x\rangle = \frac{1}{4}(||x+x||^2 - ||x-x||^2) = \frac{1}{4}(||2x||^2) = \frac{1}{4}(4||x||^2) = ||x||^2$$

, where we used property 1. We've therefore shown that $\langle x, x\rangle = ||x||^2$ for all $x \in U$. Now begins the odyssey of proving that this is, indeed, an inner product.

- **Positivity**: we've shown that $\langle x, x\rangle = ||x||^2$ and from the co-domain of our norm we know that this is always non-negative. Therefore, the property of positivity has been proved.

- **Definitiness:** again, from $\langle x, x\rangle = ||x||^2$ and from property 1 of the norm, we have that $\langle x, x\rangle = 0$ iff $x = 0$. Therefore, definitiness has been proved.

- **Symmetry:** since we're in a real vector space, the property of conjugate symmetry is simplified to $\langle x, y\rangle = \langle y, x\rangle$. To prove this, note that $\langle y, x\rangle = \frac{1}{4}(||y+x||^2 - ||y-x||^2) = \frac{1}{4}(||x+y||^2 - ||-1(x-y)||^2) = \frac{1}{4}(||x+y||^2 - |-1|||x-y||^2) = \langle x, y\rangle$.

- **Linearity in the first slot:** We shall begin by proving that $\langle x, 2y\rangle = 2\langle x, y\rangle$ for all $x, y \in U$. We have that:

$$\langle x, 2y\rangle - 2\langle x, y\rangle = \frac{1}{4}(||x+2y||^2 - ||x-2y||^2) - 2\frac{1}{4}(||x+y||^2 - ||x-y||^2) =$$
$$\frac{||x+2y||^2 - ||x-2y||^2 - 2||x+y||^2 + 2||x-y||^2}{4}$$

  Call the numerator of this fraction A. We now apply the parallelogram equality to the vectors $x+y, y$ (the reasoning being that we want to simplify the pairs of terms of A with the same sign), obtaining:

$$||x+y+y||^2 + ||x+y-y||^2 = 2(||x+y||^2 + ||y||^2) \implies ||x+2y||^2 + ||x||^2 = 2||x+y||^2 + 2||y||^2$$
$$\implies ||x+2y||^2 - 2||x+y||^2 = 2||y||^2 - ||x||^2$$

  We also apply the parallelogram equality to the vectors $x-y, -y$, obtaining:

$$||x-y-y||^2 + ||x-y-(-y)||^2 = 2(||x-y||^2 + ||-y||^2) \implies ||x-2y||^2 - 2||x-y||^2 = 2||y||^2 - ||x||^2$$

  Then, observe that we can rewrite A as:

$$A = ||x+2y||^2 - 2||x+y||^2 - (||x-2y||^2 - 2||x-y||^2) = 2||y||^2 - ||x||^2 - (2||y||^2 - ||x||^2) = 0$$

  , which in turn yields that $\langle x, 2y\rangle - 2\langle x, y\rangle = 0 \implies \langle x, 2y\rangle = 2\langle x, y\rangle$.

Let us now prove that $\langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle$ (one of the two conditions for linearity in the first slot). Firstly, we have that:

$$\langle x+y,z\rangle - (\langle x,z\rangle + \langle y,z\rangle) = \frac{1}{4}(||x+y+z||^2 - ||x+y-z||^2) - ((\frac{1}{4}(||x+z||^2 - ||x-z||^2) + \frac{1}{4}(||y+z||^2 - ||y-z||^2))$$

$$= \frac{||x+y+z||^2 - ||x+y-z||^2 - ||x+z||^2 + ||x-z||^2 - ||y+z||^2 + ||y-z||^2}{4}$$

Again, call the numerator A. A line of thought similar to the one we used for the previous step of the proof leads us to try to simplify pairs of the same sign in A. Two such pairs are $||x+y+z||^2, ||x-z||^2$ and $-||x+y-z||^2, -||x+z||^2$.

The question then is, how can we select vectors a, b such that the diagonals of the parallelogram formed by them correspond to $x+y+z, x-z$. And the answer lies in the solution of the system $a+b = x+y+z, a-b = x-z$. This will yield $a = \frac{2x+y}{2}, b = \frac{y+2z}{2}$. By the parallelogram equality:

$$||x+y+z||^2 + ||x-z||^2 = 2(||\frac{2x+y}{2}||^2 + ||\frac{y+2z}{2}||^2)$$

The same procedure yields that:

$$||x+y-z||^2 + ||x+z||^2 = 2(||\frac{2x+y}{2}||^2 + ||\frac{y-2z}{2}||^2)$$

Observe that the choice of grouping these pairs is not random: it will cause some pairs of terms to cancel out, namely:

$$A = 2(||\frac{2x+y}{2}||^2 + ||\frac{y+2z}{2}||^2) - 2(||\frac{2x+y}{2}||^2 + ||\frac{y-2z}{2}||^2) - ||y+z||^2 + ||y-z||^2$$

$$= 2||\frac{y+2z}{2}||^2 - 2||\frac{y-2z}{2}||^2 - ||y+z||^2 + ||y-z||^2 = \frac{1}{2}||y+2z||^2 - \frac{1}{2}||y-2z||^2 - ||y+z||^2 + ||y-z||^2$$

$$= \frac{2}{4}(||y+2z||^2 - ||y-2z||^2) - \frac{4}{4}(||y+z||^2 - ||y-z||^2) = 2\langle y,2z\rangle - 4\langle y,z\rangle = 4\langle y,z\rangle - 4\langle y,z\rangle = 0$$

, where in the last equality we used the previously proved fact that $\langle x,2y\rangle = 2\langle x,y\rangle$. Since $A = 0$, we've now proved linearity in the first slot.

Let us now prove that $\langle nx,y\rangle = n\langle x,y\rangle$ for every positive integer n and every $x,y \in U$. We will do this inductively:

- $n = 1$: Trivially true.
- $n = 2$: We've already proved this.
- Assume it holds for $n = k$. Then: $\langle (k+1)x,y\rangle = \langle kx+x,y\rangle = \langle kx,y\rangle + \langle x,y\rangle = k\langle x,y\rangle + \langle x,y\rangle = (k+1)\langle x,y\rangle$, where we used the induction hypothesis and the proved linearity wrt. addition in the first slot.

Continuing our gradual process of generalization, let us now prove that $m\langle \frac{1}{m}x,y\rangle = \langle x,y\rangle$ for every positive integer m and every $x,y \in U$.

We have that $\langle x,y\rangle = \langle \frac{m}{m}x,y\rangle = \langle m(\frac{x}{m}),y\rangle = m\langle \frac{x}{m},y\rangle = m\langle \frac{1}{m}x,y\rangle$, where we simply used the previously proven statement $\langle nx,y\rangle = x\langle x,y\rangle$ for n positive integer.

Let us now prove that $\langle rx,y\rangle = r\langle x,y\rangle$ for r rational number, $x,y \in U$. Since r is rational, we have that $r = \frac{\kappa}{\lambda}$ for κ, λ integers (assume, for now, positive). Then, by using the two statements proved above:

$$\langle rx,y\rangle = \langle \frac{\kappa}{\lambda}x,y\rangle = \kappa\langle \frac{1}{\lambda}x,y\rangle = \kappa\frac{\lambda}{\lambda}\langle \frac{1}{\lambda}x,y\rangle = \frac{\kappa}{\lambda}\langle x,y\rangle = r\langle x,y\rangle$$

, where we directly used the two statements proved above for integers. Furthermore, we note that $\langle -x, y\rangle = \frac{1}{4}(||-x+y||^2 - ||-x-y||^2) = \frac{1}{4}(||x-y||^2 - ||x+y||^2) = -\langle x, y\rangle$, so it holds for negative rational numbers as well.

We now proceed to prove that the Cauchy-Schwarz inequality holds. We begin by applying the triangle inequality for any $x, y \in U$, which we know holds (property 3). First, we note that due to linearity wrt. addition in the first slot and due to the symmetry of the inner product, it's also true that $\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle$. Then:

$$||x+y|| \leq ||x|| + ||y|| \implies ||x+y||^2 \leq (||x|| + ||y||)^2 \implies \langle x+y, x+y\rangle \leq ||x||^2 + 2||x||||y|| + ||y||^2$$

$$\implies \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \leq ||x||^2 + 2||x||||y|| + ||y||^2 \implies \langle x, y\rangle \leq ||x||||y||$$

Then:

$$||x-y|| \leq ||x|| + ||-y|| \implies ||x-y||^2 \leq (||x|| + ||y||)^2 \implies \langle x-y, x-y\rangle \leq ||x||^2 + 2||x||||y|| + ||y||^2$$

$$\implies \langle x, x\rangle - \langle x, y\rangle - \langle y, x\rangle + \langle y, y\rangle \leq ||x||^2 + 2||x||||y|| + ||y||^2 \implies -2\langle x, y\rangle \leq 2||x||||y||$$

$$\implies \langle x, y\rangle \geq -||x||||y||$$

These two facts yield $|\langle x, y\rangle| \leq ||x||||y||$.

Now we prove that for $c \in \mathbf{R}$, r rational number and $x, y \in U$,

$$|c\langle x, y\rangle - \langle cx, y\rangle| = |(c-r)\langle x, y\rangle - \langle (c-r)x, y\rangle| \leq 2|c-r|||x||||y||$$

Firstly:

$$|(c-r)\langle x, y\rangle - \langle (c-r)x, y\rangle| = |c\langle x, y\rangle - r\langle x, y\rangle - \langle cx - rx, y\rangle| = |c\langle x, y\rangle - r\langle x, y\rangle - \langle cx, y\rangle - \langle -rx, y\rangle|$$

$$= |c\langle x, y\rangle - r\langle x, y\rangle - \langle cx, y\rangle + r\langle x, y\rangle| = |c\langle x, y\rangle - \langle cx, y\rangle|$$

, where we used the previously proven statement about rational numbers. Now, by the triangle inequality (on real numbers) and the Cauchy-Schwarz inequality:

$$|(c-r)\langle x, y\rangle - \langle (c-r)x, y\rangle| \leq |c-r||\langle x, y\rangle| + |\langle (c-r)x, y\rangle| \leq |c-r|||x||||y|| + ||(c-r)x||||y||$$

$$= 2|c-r|||x||||y||$$

, which completes the proof of the statement.

Having proved this, consider an arbitrary $c \in \mathbf{R}$. We know that we can approach "arbitrarily close" to c with a sequence of rational numbers, i.e. $|c-r|$ can be made arbitrarily small. Therefore, by the previous inequality, we can make $|c\langle x, y\rangle - \langle cx, y\rangle|$ arbitrarily small as well. This means, then, that $c\langle x, y\rangle = \langle cx, y\rangle$, which establishes homogeneity wrt. the first slot of the "aspiring" inner product.

Having proved these properties, we conclude that $\langle , \rangle$ is indeed an inner product on U and is associated with the given norm. This concludes the proof for real vector spaces.
We will now prove another useful lemma, namely that if V is a complex inner product space with an inner product $\langle ., .\rangle$, and $[., .]$ is the real-valued function such that $[x, y] = Re\{\langle x, y\rangle\}$ for all $x, y \in V$, then $[., .]$ is an inner product over V, where V is regarded as a vector space over $\mathbf{R}$. Additionally, we'll show that $[x, ix] = 0$ for all $x \in V$. We examine the properties necessary and sufficient for a function to be considered an inner product one by one:

- **Positivity**: For $x \in V$, we have that $[x, x] = Re\{\langle x, x\rangle\} \geq 0$, since it is true that $\langle x, x\rangle \geq 0$ from the definition of inner product applied on$\langle ., .\rangle$.
- **Definitiness**: For $x \in V$, we are interested in when $[x, x]$ is zero. We have, then, that: $[x, x] = 0 \iff Re\{\langle x, x\rangle\} = 0$, which since $\langle x, x\rangle \in \mathbf{R}$ is equivalent to $\langle x, x\rangle = 0$, which happens iff $x = 0$, again from the definition of inner product applied on $\langle ., .\rangle$. Therefore $[x, x] = 0 \iff x = 0$.

- **Linearity in the first slot**: Firstly, for $x, y, z \in V$: $[x+y, z] = Re\{\langle x+y, z\rangle\} = Re\{\langle x, z\rangle + \langle y, z\rangle\} = Re\{\langle x, z\rangle\} + Re\{\langle y, z\rangle\} = [x, z] + [y, z]$, where we used the additivity of the real part as well as the additivity in the first slot of $\langle ., .\rangle$. Therefore, additivity in the first slot holds for $[., .]$.

  Secondly, since we are considering V as a vector space over $\mathbf{R}$, we need to show that $[\alpha x, y] = \alpha[x, y]$ for $\alpha \in \mathbf{R}$ and $x, y \in V$. We have that $[\alpha x, y] = Re\{\langle \alpha x, y\rangle\} = Re\{\alpha\langle x, y\rangle\} = \alpha Re\{\langle x, y\rangle\} = \alpha[x, y]$, where we used the homogeneity in the first slot of $\langle ., .\rangle$ and the fact that $\alpha \in \mathbf{R}$, which means that $Re\{\alpha c\} = \alpha Re\{c\}$ for any $c \in \mathbf{C}$.

- **(Conjugate) Symmetry**: Since we're discussing a real inner product, the symmetry property states that $[y, x] = [x, y]$ for $x, y \in V$. To prove this, we have that $[y, x] = Re\{\langle y, x\rangle\} = Re\{\overline{\langle x, y\rangle}\} = Re\{\langle x, y\rangle\} = [x, y]$, where we used the conjugate symmetry property of $\langle ., .\rangle$ as well as the fact that $Re\{\bar{c}\} = Re\{c\}$ for all $c \in \mathbf{C}$.

  Therefore, $[., .]$ is indeed an inner product over V, where V is regarded as a vector space over $\mathbf{R}$. We also have that:
  $$[x, ix] = Re\{\langle x, ix\rangle\} = Re\{\bar{i}\langle x, x\rangle\} = Re\{-i\langle x, x\rangle\} = 0$$
  , since $\langle x, x\rangle \in \mathbf{R}$ for all $x \in V$, which implies that $-i\langle x, x\rangle$ is an imaginary number, i.e. real part zero.

Continuing the process of proving the initial statement for complex vector spaces, we'll need the following lemma as well:
Let V be a vector space over $\mathbf{C}$ and suppose that $[., .]$ is a real inner product on V, where V is regarded as a vector space over $\mathbf{R}$, such that $[x, ix] = 0$ for all $x \in V$. Let $\langle ., .\rangle$ be the complex-valued function defined by
$$\langle x, y\rangle = [x, y] + i[x, iy]$$
for $x, y \in V$. Then, $\langle ., .\rangle$ is a complex inner product on V. Once again, we examine the properties that define an inner product one by one:

- **Positivity**: We have that $\langle x, x\rangle = [x, x] + i[x, ix] = [x, x] + i \cdot 0 = [x, y] \geq 0$, since $[., .]$ is an inner product on V.

- **Definiteness**: We have that $\langle x, x\rangle = 0 \iff [x, x] + i[x, ix] = 0 \iff [x, x] = 0$, which happens if and only if $x = 0$ since $[., .]$ is an inner product on V.

- **Linearity in the first slot**: Firstly, for $x, y, z \in V$:
  $$\langle x+y, z\rangle = [x+y, z] + i[x+y, iz] = [x, z] + [y, z] + i([x, iz] + [y, iz]) =$$
  $$[x, z] + i[x, iz] + [y, z] + i[y, iz] = \langle x, z\rangle + \langle y, z\rangle$$
  , where we've used the linearity of $[., .]$ as an inner product. Secondly, let $\alpha = b + ci \in \mathbf{C}$ be a scalar. Then:
  $$\langle \alpha x, y\rangle = \langle (b+ci)x, y\rangle = [(b+ci)x, y] + i[(b+ci)x, iy] = [bx + cix, y] + i[bx + cix, iy]$$
  $$= [bx, y] + [cix, y] + i[bx, iy] + i[cix, iy] = b[x, y] + c[ix, y] + bi[x, iy] + ci[ix, iy] =$$
  $$b([x, y] + i[x, iy]) + ci([ix, iy] - i[ix, y]) = b\langle x, y\rangle + ci([ix, iy] - i[ix, y])$$

  Observe that from the equation above it becomes clear that if we show that $\langle x, y\rangle = [ix, iy] - i[ix, y]$, then we'll also have shown homogeneity in the first slot. To show this, we have that, for any $x, y \in V$:
  $$[x+y, i(x+y)] = 0 \implies [x, ix] + [x, iy] + [y, ix] + [y, iy] = 0 \implies [x, iy] + [y, ix] = 0$$
  $$\implies [x, iy] = -[y, ix]$$
  , where we've used the fact that the inner product $[., .]$ is additive in the first and second slots (second slot additivity is provable, as is done in the book). Therefore: $[ix, iy] = -[y, i(ix)] = -[y, -x] =$

$[y,x] = [x,y]$, by using the symmetry and homogeneity of $[.,]$. Furthermore, $[ix,y] = [ix,i(-iy)] = -[-iy,i(ix)] = [iy,-x] = -[x,iy]$, once again using the symmetry and homogeneity of $[.,.]$.

By these two facts, $[ix,iy] - i[ix,y] = [x,y] - i(-[x,iy]) = [x,y] + i[x,iy] = \langle x,y\rangle$. As observed above, this essentially implies that $\langle \alpha x,y\rangle = \alpha\langle x,y\rangle$ for $\alpha \in \mathbf{C}$. Therefore, we've now shown linearity in the first slot for $\langle,\rangle$.

- **Conjugate symmetry**: We have that $\overline{\langle y,x\rangle} = \overline{[y,x] + i[y,ix]} = \overline{[y,x]} + \overline{i[y,ix]} = [y,x] - i[y,ix] = [x,y] - i[ix,y] = [x,y] - i(-[x,iy]) = [x,y] + i[x,iy] = \langle x,y\rangle$, where we applied the symmetry of $[.,.]$ (real inner product) as well as the previously proven fact $[ix,y] = -[x,iy]$. We've therefore shown that conjugate symmetry holds for $\langle.,.\rangle$.

Therefore, $\langle.,.\rangle$ is indeed a complex inner product on V.
Finally, let's use all of the lemmas we've proved to show that a norm satisfying the parallelogram equality comes from an inner product in complex vector spaces as well. Suppose then V is a complex vector space and $||\cdot||$ is a norm as defined in the exercise. Then:

- Let $[x,y] = \frac{1}{4}(||x+y||^2 - ||x-y||^2)$. We proved that if we consider V as a vector space over $\mathbf{R}$, $[.,.]$ is an inner product for which $\langle x,x\rangle = ||x||^2$ for all $x \in V$. It is furthermore the case that:

$$[x,ix] = \frac{1}{4}(||x+ix||^2 - ||x-ix||^2) = \frac{1}{4}(|1+i||x||^2 - |1-i||x||^2) = \frac{1}{4}(\sqrt{2}||x||^2 - \sqrt{2}||x||^2) = 0$$

- Now, let $\langle x,y\rangle = [x,y] + i[x,iy]$. Since $[.,.]$ is an inner product on V regarded as a vector space over $\mathbf{R}$ and has the property that $[x,ix] = 0$, we know from the previously proved lemma that $\langle.,.\rangle$ is a complex inner product on V.

- Lastly, for $x \in V$: $\langle x,x\rangle = [x,x] + i[x,ix] = ||x||^2$. We conclude, therefore, that $\langle.,.\rangle$ is an inner product on V associated with the given norm, and thus we — finally — conclude the proof.

Exercise 26
Suppose f,g are differentiable functions from $\mathbf{R}$ to $\mathbf{R}^n$.
(a) Show that

$$\langle f(t),g(t)\rangle' = \langle f'(t),g(t)\rangle + \langle f(t),g'(t)\rangle$$

(b) Suppose $c > 0$ and $||f(t)|| = c$ for every $t \in \mathbf{R}$. Show that $\langle f'(t),f(t)\rangle = 0$ for every $t \in \mathbf{R}$.
(c) Interpret the result in part (b) geometrically in terms of the tangent vector to a curve lying on a sphere in $\mathbf{R}^n$ centered at the origin.

Solution. (a)
By using the definition of the derivative, we have that:

$$\langle f(t),g(t)\rangle' = \lim_{h\to 0}\left(\frac{\langle f(t+h),g(t+h)\rangle - \langle f(t),g(t)\rangle}{h}\right) =$$

$$= \lim_{h\to 0}\left(\frac{\langle f(t+h),g(t+h)\rangle - \langle f(t+h),g(t)\rangle + \langle f(t+h),g(t)\rangle - \langle f(t),g(t)\rangle}{h}\right) =$$

$$\lim_{h\to 0}\left(\frac{\langle f(t+h),g(t+h)-g(t)\rangle + \langle f(t+h)-f(t),g(t)\rangle}{h}\right) =$$

$$\lim_{h\to 0}\left(\langle f(t+h),\frac{g(t+h)-g(t)}{h}\rangle + \langle\frac{f(t+h)-f(t)}{h},g(t)\rangle\right)$$

, where we used the first and second slot linearity of the inner product, as well as its symmetry as a real inner product

Observe now that if it holds that $\lim_{x\to x_0}(\langle f_1(x), f_2(x)\rangle) = \langle \lim_{x\to x_0}(f_1(x)), \lim_{x\to x_0}(f_2(x))\rangle$, for any x_0 and any f_1, f_2 functions whose limits at x_0 are well defined, then the equality above essentially yields what we require. This is because:

$$\lim_{h\to 0}(f(t+h)) = f(t), \lim_{h\to 0}\left(\frac{g(t+h)-g(t)}{h}\right) = g'(t), \lim_{h\to 0}\left(\frac{f(t+h)-f(t)}{h}\right) = f'(t), \lim_{h\to 0}(g(t)) = g(t)$$

In fact, if we can show that $\langle \cdot, \cdot \rangle$ is continuous, then the composition rule for limits in combination with the continuity of our f, g will guarantee the equality we required above (and which completes the proof). In other words, we need to show that the function $F : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}, F(x_1, x_2) = \langle x_1, x_2\rangle$ is continuous. This is equivalent to showing that:

$$\forall a, b \in \mathbf{R}^n, \lim_{x_1\to a, x_2\to b}(\langle x_1, x_2\rangle) = \langle a, b\rangle$$

, which can be shown by proving that for any $\epsilon > 0$, we can choose $\delta > 0$ such that whenever $||x_1 - a|| \leq \delta, ||x_2 - a|| \leq \delta$ it is true that $|\langle x_1, x_2\rangle - \langle a, b\rangle| \leq \epsilon$. Let us pick therefore any $\epsilon > 0$. We begin by observing that:

$$|\langle x_1, x_2\rangle - \langle a, b\rangle| = |\langle x_1, x_2\rangle - \langle a, x_2\rangle + \langle a, x_2\rangle - \langle a, b\rangle| = |\langle x_1 - a, x_2\rangle + \langle a, x_2 - b\rangle| \leq |\langle x_1 - a, x_2\rangle| + |\langle a, x_2 - b\rangle|$$

$$\leq ||x_1 - a|| \cdot ||x_2|| + ||a|| \cdot ||x_2 - b||$$

, where at the last step we used the Cauchy-Schwarz inequality. Now pick $\delta = \min \frac{\epsilon}{2}, \frac{\epsilon}{2(||b||+\epsilon)}, \frac{\epsilon}{2(||a||+\epsilon)}$. Observe that for this choice of δ, whenever $||x_2 - b|| \leq \delta, ||x_1 - a|| \leq \delta$ we have that:

$$||a|| \cdot ||x_2 - b|| \leq ||a|| \frac{\epsilon}{2(\epsilon + ||a||)} \leq \frac{\epsilon}{2}$$

Furthermore $||x_2|| - ||b|| \leq ||x_2 - b|| \leq \epsilon$ (the first inequality follows easily from squaring and using the Cauchy-Schwarz inequality), which means that $||x_2|| \leq \epsilon + ||b||$. Then:

$$||x_1 - a|| \cdot ||x_2|| \leq \frac{\epsilon}{2(||b|| + \epsilon)} \cdot (\epsilon + ||b||) = \frac{\epsilon}{2}$$

. Putting it all together, we conclude that for the above choice of δ it holds that

$$|\langle x_1, x_2\rangle - \langle a, b\rangle| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

.

We therefore proved that the inner product function is continuous, which combined with the observations we made prior to this proof, yields the result $\langle f(t), g(t)\rangle' = \langle f'(t), g(t)\rangle + \langle f(t), g'(t)\rangle$.

(b)

Observe that $\langle f(t), f(t)\rangle' = \langle f'(t), f(t)\rangle + \langle f(t), f'(t)\rangle \implies (||f(t)||^2)' = 2\langle f(t), f'(t)\rangle$. If $||f(t)|| = c$ for every $t \in \mathbf{R}$, then $||f(t)||^2 = c^2 \implies (||f(t)||^2)' = 0$ for all $t \in \mathbf{R}$.

By the previous observation, we have then that for every $t \in \mathbf{R}$, $\langle f(t), f'(t)\rangle = 0$.

(c)

If $||f(t)|| = c$ for all $t \in \mathbf{R}$, every point of the (parameterized) curve of f lies on a sphere of radius c centered at the origin. At any point $f(t)$ of this sphere, the tangent plane comprises all vectors that are orthogonal to the vector $f(t)$. Since the curve lies entirely on the sphere's surface, $f'(t)$ will belong to the tangent plane of the sphere at point $f(t)$, and will thus be orthogonal to $f(t)$, as we showed in (b).

6.B Orthonormal Bases

Exercise 1
(a) Suppose $\theta \in \mathbf{R}$. Show that $(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)$ and $(\cos\theta, \sin\theta), (\sin\theta, -\cos\theta)$ are orthonormal bases of $\mathbf{R}^2$.
(b) Show that each orthonormal basis of $\mathbf{R}^2$ is of the form given by one of the two possibilities of part (a).

Solution.
(a)
For the first list, we have that $||(\cos\theta, \sin\theta)|| = \sqrt{\cos^2\theta + \sin^2\theta} = 1$ and
$||(-\sin\theta, \cos\theta)|| = \sqrt{(-\sin\theta)^2 + \cos^2\theta} = 1$, by the well known trigonometric identity. Furthermore, $\langle(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)\rangle = -\sin\theta\cos\theta + \sin\theta\cos\theta = 0$, so the list is indeed orthonormal (orthogonal vectors with norm 1). The proof is almost identical for the second list given in the exercise.
(b)
Let v_1, v_2 be an orthonormal basis of $\mathbf{R}^2$. First of all, observe that if we define as θ the angle of a vector v of length $\alpha > 0$ with the $x'x$ axis (the angle defined by the mathematical positive rotation of the $x'x$ axis), we can express v as $v = (\alpha\cos\theta, \alpha, \sin\theta)$. Therefore, let $v_1 = (\alpha_1\cos\theta_1, \alpha_1\sin\theta_1), v_2 = (\alpha_2\cos\theta_2, \alpha_2\sin\theta_2)$ be the vectors of our orthonormal basis expressed in this formulation. Since the basis is orthonormal, the vectors must each have norm 1, so $\alpha_1 = \alpha_2 = 1$. Furthermore, it must hold that:

$$\langle v_1, v_2\rangle = 0 \implies \cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2 = 0 \implies \cos\theta_1\cos\theta_2 = -\sin\theta_1\sin\theta_2$$

If $\cos\theta_1 = 0$, then $\sin\theta_1 = \pm 1$ and by the above equality we have that $\sin\theta_2 = 0$, therefore $\cos\theta_2 = \pm 1$. In any case though, it holds that: $\sin\theta_2 = \cos\theta_1$ and $\sin\theta_2 = -\cos\theta_1$ (since they're both zero) and either that $\cos\theta_2 = \sin\theta_1$ or that $\cos\theta_2 = -\sin\theta_1$. It's then clear that the orthonormal basis has one of the forms of (a).
If $\cos\theta_1 \neq 0$, then by the above equality, $\cos\theta_2 = -\frac{\sin\theta_1\sin\theta_2}{\cos\theta_1}$. By plugging this in the basic trigonometric identity for θ_2 we obtain that:

$$\cos^2\theta_2 + \sin^2\theta_2 = 1 \implies \sin^2\theta_2 + \frac{\sin^2\theta_1\sin^2\theta_2}{\cos^2\theta_1} = 1 \implies \sin^2\theta_2\cos^2\theta_1 + \sin^2\theta_1\sin^2\theta_2 = \cos^2\theta_1$$

$$\implies \sin^2\theta_2(\cos^2\theta_1 + \sin^2\theta_1) = \cos^2\theta_1 \implies \sin^2\theta_2 = \cos^2\theta_1$$

This implies that either $\sin\theta_2 = \cos\theta_1$ or that $\sin\theta_2 = -\cos\theta_1$. In the first case, we have that $\cos\theta_2 = -\frac{\sin\theta_1\sin\theta_2}{\cos\theta_1} = -\frac{\sin\theta_1\cos\theta_1}{\cos\theta_1} = -\sin\theta_1$. Therefore, $v_2 = (-\sin\theta_1, \cos\theta_1)$ and $v_1 = (\cos\theta_1, \sin\theta_1)$, which is of the first possibility proposed in (a). In the second case, we have that $\cos\theta_2 = -\frac{\sin\theta_1\sin\theta_2}{\cos\theta_1} = -\frac{\sin\theta_1(-\cos\theta_1)}{\cos\theta_1} = \sin\theta_1$. Therefore, $v_2 = (\sin\theta_1, -\cos\theta_1)$ and $v_1 = (\cos\theta_1, \sin\theta_1)$, which is of the second possibility proposed in(a).
In any case, since v_1, v_2 was selected as an arbitrary orthonormal basis of $\mathbf{R}^2$, we have proved that each orthonormal basis of $\mathbf{R}^2$ is of the form proposed in part (a).

Exercise 2
Suppose $e_1, e_2, \dots, e_m$ is an orthonormal list of vectors in V. Let $v \in V$. Prove that

$$||v||^2 = |\langle v, e_1\rangle|^2 + \dots + |\langle v, e_m\rangle|^2$$

if and only if $v \in \text{span}(e_1, \dots, e_m)$.

Solution.
$\Longleftarrow$: If $v \in \text{span}(e_1, \dots, e_m)$, and $e_1, \dots, e_m$ is an orthonormal list, we already know that $||v||^2 = |\langle v, e_1\rangle|^2 + \dots + |\langle v, e_m\rangle|^2$.
$\Longrightarrow$: If $||v||^2 = |\langle v, e_1\rangle|^2 + \dots + |\langle v, e_m\rangle|^2$, we have that:

$$\langle v - \sum_i \langle v, e_i\rangle e_i, v - \sum_j \langle v, e_j\rangle e_j\rangle = \langle v, v\rangle - \langle v, \sum_j \langle v, e_j\rangle e_j\rangle - \langle \sum_i \langle v, e_i\rangle e_i, v\rangle + \langle \sum_i \langle v, e_i\rangle e_i, \sum_j \langle v, e_j\rangle e_j\rangle = A$$

, where we used the first and second slot additivity of the inner product. Continuing, we have that:

$$A = ||v||^2 - \sum_i \overline{\langle v, e_i\rangle}\langle v, e_i\rangle - \sum_i \langle v, e_i\rangle\langle e_i, v\rangle + \sum_i \langle v, e_i\rangle\langle e_i, \sum_j \langle v, e_j\rangle e_j\rangle$$

, where we again used first and second slot additivity as well as first slot homogeneity and second slot conjugate homogeneity of the inner product. Then:

$$A = ||v||^2 - \sum_i |\langle v, e_i\rangle|^2 - \sum_i \langle v, e_i\rangle\langle e_i, v\rangle + \sum_i \langle v, e_i\rangle(\sum_j \overline{\langle v, e_j\rangle}\langle e_i, e_j\rangle) =$$
$$= ||v||^2 - \sum_i |\langle v, e_i\rangle|^2 - \sum_i \langle v, e_i\rangle\overline{\langle v, e_i\rangle} + \sum_i \langle v, e_i\rangle(\sum_j \overline{\langle v, e_j\rangle}\langle e_i, e_j\rangle)$$

, where we used the conjugate symmetry of the inner product. Now, for the last term of this sum observe that due to e_i being orthonormal, $\langle e_i, e_j\rangle = 0$ except when $i = j$, in which case $\langle e_i, e_j\rangle = 1$. Therefore:

$$A = ||v||^2 - 2\sum_i |\langle v, e_i\rangle|^2 + \sum_i \langle v, e_i\rangle\overline{\langle v, e_i\rangle} = \sum_i |\langle v, e_i\rangle|^2 - 2\sum_i |\langle v, e_i\rangle|^2 + \sum_i |\langle v, e_i\rangle|^2 = 0$$

, where we used the assumption of the " $\Longrightarrow$ " direction. Since $A = \langle v - \sum_i \langle v, e_i\rangle e_i, v - \sum_j \langle v, e_j\rangle e_j\rangle$, and since $\langle x, x\rangle = 0$ iff $x = 0$, we conclude that $v = \sum_i \langle v, e_i\rangle e_i$, hence $v \in \text{span}(e_1, \dots, e_m)$, completing the proof.

Exercise 4
Suppose n is a positive integer. Prove that

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos(x)}{\sqrt{\pi}}, \frac{\cos(2x)}{\sqrt{\pi}}, \dots, \frac{\cos(nx)}{\sqrt{\pi}}, \frac{\sin(x)}{\sqrt{\pi}}, \frac{\sin(2x)}{\sqrt{\pi}}, \dots, \frac{\sin(nx)}{\sqrt{\pi}}$$

is an orthonormal list of vectors in $C[-\pi, \pi]$, the vector space of continuous real-valued functions on $[-\pi, \pi]$ with the inner product

$$\langle f, g\rangle = \int_{-\pi}^{\pi} f(x)g(x)dx$$

Solution.
We begin by examining the norms of each vector.

- $||\frac{1}{\sqrt{2\pi}}|| = \int_{-\pi}^{\pi} \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}dx = \left(\frac{x}{2\pi}\right)\Big|_{-\pi}^{\pi} = \frac{\pi}{2\pi} - (-\frac{\pi}{2\pi}) = 1$
- $||\frac{\cos(nx)}{\sqrt{pi}}|| = \int_{-\pi}^{\pi} \frac{\cos(nx)}{\sqrt{pi}}\frac{\cos(nx)}{\sqrt{pi}}dx = \left(\frac{1}{2\pi}(x + \frac{\sin(2nx)}{2n})\right)\Big|_{-\pi}^{\pi} = \frac{1}{2\pi}(\pi + 0 - (-\pi + 0)) = 1$, where we omitted several steps in the calculation of the — indefinite —integral of $\cos^2(nx)$ (trig. identities etc.) since they're more relevant to a calculus course. Note that we this is true for any positive integer n.

- $||\frac{\sin(nx)}{\sqrt{pi}}|| = \int_{-\pi}^{\pi} \frac{\sin(nx)}{\sqrt{pi}} \frac{\sin(nx)}{\sqrt{pi}} dx = \left(-\frac{\sin(2nx)-2nx}{4\pi n}\right)\Big|_{-\pi}^{\pi} = (-\frac{0-2n\pi}{4\pi n}) - (-\frac{0+2n\pi}{4\pi n}) = 1$, again by omitting several steps, and again true for any positive n.

Therefore every vector of the list has norm 1. Now we need to show that $\langle f_1, f_2 \rangle = 0$ whenever f_1, f_2 are vectors of the list which are not equal.

- If $f_1 = \frac{1}{\sqrt{2\pi}}$, and $f_2 = \frac{\cos(nx)}{\sqrt{\pi}}$, then $\int_{-\pi}^{\pi} \frac{1}{\sqrt{2\pi}} \frac{\cos(nx)}{\sqrt{\pi}} dx = \left(\frac{1}{\pi\sqrt{2}} \frac{\sin(nx)}{n}\right)\Big|_{-\pi}^{\pi} = 0 - 0$, for any positive integer n.
- If $f_1 = \frac{1}{\sqrt{2\pi}}$, and $f_2 = \frac{\sin(nx)}{\sqrt{\pi}}$, then $\int_{-\pi}^{\pi} \frac{1}{\sqrt{2\pi}} \frac{\sin(nx)}{\sqrt{\pi}} dx = 0$ because sin is an odd function.
- If $f_1 = \frac{\cos(n_1 x)}{\sqrt{\pi}}$ and $f_2 = \frac{\sin(n_2 x)}{\sqrt{\pi}}$, then observe that their product is an odd function because cos is even and sin is odd, so its integral from $-\pi$ to π will be 0.
- If f_1, f_2 are both cosines or both sines, we can show that they are orthogonal by computing the respective integrals. Again, however, this is rather tedious and has more to do with calculus practice, so we'll omit it.

Therefore, we've also shown that all of the vectors in the list are orthogonal, hence proving the orthonormality of the list.

Exercise 9
What happens if the Gram-Schmidt Procedure is applied to a list of vectors that is not linearly independent?

Solution.
If the first vector of this list is the zero vector, then the Gram-Schmidt procedure will try to divide with zero (the vector's norm). Otherwise, by the linear dependence lemma, there exists some $j > 1$ such that $v_j \in \text{span}(v_1, \ldots, v_{j-1})$, with $v_1, \ldots, v_{j-1}$ being linearly independent (this can be obtained by picking the minimum j for which the lemma holds).
Therefore, the Gram-Schmidt Procedure will work normally until v_j, yielding the orthonormal vectors $e_1, \ldots, e_{j-1}$. At this point, observe that $v_j \in \text{span}(v_1, \ldots, v_{j-1})$ and that from the Procedure we know that $\text{span}(v_1, \ldots, v_{j-1}) = \text{span}(e_1, \ldots, e_{j-1})$, thus $v_j \in \text{span}(e_1, \ldots, e_{j-1})$ and since this is an orthonormal list, we know that $v_j = \langle v_j, e_1 \rangle e_1 + \ldots + \langle v_j, e_{j-1} \rangle e_{j-1}$. Clearly, this implies that the Gram-Schmidt procedure would again try to divide by zero (the magnitude of the difference of the two sides of this equation).

Exercise 15
Suppose $C_{\mathbf{R}}([-1,1])$ is the vector space of continuous real-valued functions on the interval $[-1,1]$ with inner product given by

$$\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)dx$$

for $f, g \in C_{\mathbf{R}}([-1,1])$. Let ϕ be the linear functional on $C_{\mathbf{R}}([-1,1])$ defined by $\phi(f) = f(0)$. Show that there does not exist $g \in C_{\mathbf{R}}([-1,1])$ such that

$$\phi(f) = \langle f, g \rangle$$

for every $f \in C_{\mathbf{R}}([-1,1])$.

Solution.
Suppose that such a g existed.

- First of all, observe that $g(x)$ cannot be the zero function, because then $\langle f, g \rangle = 0$ for every f, which would mean that it could not, for example, express the value $f(0)$ for the function $f(x) = 1$.

- Let f_1 be a function in our vector space such that $f_1(0) = 0$ and $f_1(x) > 0$ for all $x \neq 0$ (such a function can easily be constructed, e.g. a second-degree polynomial). If it was the case that $g(x) \geq 0$ for all $x \in (-1, 1)$, then $\langle f_1, g\rangle > 0$, since the function $f(x)g(x)$ being integrated would be continuous (product of continuous functions in $C_{\mathbf{R}}([-1, 1])$) and non-negative everywhere, and since $f_1 \neq 0, g \neq 0$. But, $f(0) = 0$ so $\langle f_1, g\rangle \neq f_1(0)$. Therefore, there exists some $x_1 \in (-1, 1)$ such that $g(x_1) < 0$.

- Let f_2 be a function in our vector space such that $f_2(0) = 0$ and $f_2(x) < 0$ for all $x \neq 0$ (again, it's easy to construct such a function). By the same argument as before, if it was the case that $g(x) \leq 0$ everywhere, then $\langle f_2, g\rangle < 0$ and $f_2(0) = 0$, a contradiction. Therefore, there exists some $x_2 \in (-1, 1)$ such that $g(x_2) > 0$.

Without loss of generality, $x_1 < x_2$. Since g is continuous, we can apply the intermediate value theorem in $[x_1, x_2]$ and obtain that $g(\xi) = 0$ for at least one $\xi \in [x_1, x_2]$. Furthermore, by the same theorem, $g(x) \leq 0$ in a neighborhood $[x_3, \xi]$ and $g(x) \geq 0$ for a neighborhood $[\xi, x_4]$. Because g is not zero everywhere, it is also the case that we can pick x_3, x_4 such that even if g is zero in a neighborhood around ξ, there do exist neighborhoods around $[x_3, a], [b, x_4]$ where g is strictly negative. Let us then consider a function f such that:

$$f(x) = \begin{cases} -x + x_3, & x_3 \leq x \leq \frac{\xi + x_3}{2} \\ x - \xi, & \frac{\xi + x_3}{2} < x \leq \frac{\xi + x_4}{2} \\ -x + x_4, & \frac{\xi + x_4}{2} < x \leq x_4 \\ 0, & \text{elsewhere} \end{cases}$$

f is certainly continuous (it's essentially a "triangular" pulse), and we can see that $f(x)g(x) = 0$ for $x \notin [x_3, x_4]$, $f(x)g(x) \geq 0$ for $x \in [x_3, x_4]$, since we designed f such that it has the same sign as g in $[x_3, x_4]$. Furthermore, the product is not always 0 due to g, f not being always zero. Thus $\langle f, g\rangle > 0$. If $\xi = 0$, then $f(0) = f(\xi) = 0$. If $\xi \neq 0$, then we can choose x_3, x_4 arbitrarily close to it, and thus we can safely assume that $0 \notin (x_3, x_4)$, which means that $f(0) = 0$ (since 0 is outside the interval where f is non-zero). This means that $\langle f, g\rangle \neq f(0)$, a contradiction.
Therefore, such a g cannot exist.

Exercise 17
For $u \in V$, let Φu denote the linear functional on V defined by

$$(\Phi u)(v) = \langle v, u\rangle$$

for $v \in V$.
(a) Show that if $\mathbf{F} = \mathbf{R}$, then Φ is a linear map from V to V'.
(b) Show that if $\mathbf{F} = \mathbf{C}$ and $V \neq 0$, then Φ is not a linear map.
(c) Show that Φ is injective.
(d) Suppose $\mathbf{F} = \mathbf{R}$ and V is finite-dimensional. Use parts (a) and (c) and a dimension-counting argument (but without using 6.42) to show that Φ is an isomorphism from V to V'.

Solution.
(a) For $u_1, u_2 \in V$, we have that $\Phi(u_1 + u_2) = f$, where f is a function from V to $\mathbf{R}$ such that $f(v) = \langle v, u_1 + u_2\rangle = \langle v, u_1\rangle + \langle v, u_2\rangle$, due to additivity of the inner product. It's also the case that $\Phi(u_1) = f_1, \Phi(u_2) = f_2$ such that $f_1(v) = \langle v, u_1\rangle, f_2(v) = \langle v, u_2\rangle$, which leads to the conclusion that for every $v \in V, f(v) = f_1(v) + f_2(v)$, therefore $\Phi(u_1 + u_2) = \Phi(u_1) + \Phi(u_2)$.
Furthermore, for $\lambda \in \mathbf{R}$ and $u \in V$, $\Phi(\lambda u) = f$, where f is a function from V to $\mathbf{R}$ such that for $v \in V, f(v) = \langle v, \lambda u\rangle = \lambda\langle v, u\rangle$, where we used the second-slot homogeneity of the inner product (which holds due to $\mathbf{F} = \mathbf{R}$). This means that $\Phi(\lambda u) = \lambda\Phi(u)$.
We have thus proved linearity and homogeneity for Φ. Obviously, the domain of Φ is V, whereas $\Phi(u) = f$, such that $f(v) = \langle v, u\rangle$, and because the inner product is linear in its first slot, f is linear, thus the codomain of Φ is indeed $L(V, \mathbf{F}) = V'$.

(b) If $\mathbf{F} = \mathbf{C}$, $V \neq 0$, then observe that for $\lambda \in \mathbf{C}$, $u \in V, u \neq 0$ (such a u exists because $V \neq 0$), $\Phi(\lambda u) = f$, such that $f(v) = \langle v, \lambda u \rangle = \overline{\lambda}\langle v, u \rangle$. Clearly, if $\lambda \notin \mathbf{R}$ this is not equal to $\Phi(u)$, therefore homogeneity does not hold and Φ cannot be linear.
(c) We need to examine when is it the case that $\Phi(u)$ is the zero function, i.e. when is it true that $\langle v, u \rangle = 0$ for every $v \in V$. This is equivalent to u being orthogonal to every vector $v \in V$, which happens iff $u = 0$, therefore $\Phi(u) \iff u = 0$, i.e. Φ is injective.
(d) From (c), it is the case that $\dim(\text{null}\Phi) = 0$, therefore by the Fundamental Theorem of Linear Maps, $\dim V = \dim(\text{null}\Phi) + \dim(\text{range}\Phi) \implies \dim V = \dim(\text{range}\Phi)$, and we know that $\dim V = \dim V'$, therefore Φ is surjective, and since it's also injective, it's invertible, i.e. it's an isomorphism between V and V'.

6.C Orthogonal Complements and Minimization Problems

Exercise 1
Suppose $v_1, \ldots, v_m \in V$. Prove that

$$\{v_1, \ldots, v_m\}^{\perp} = (\text{span}(v_1, \ldots, v_m))^{\perp}$$

Solution.
Let $S_1 = \{v_1, \ldots, v_m\}^{\perp}, S_2 = (\text{span}(v_1, \ldots, v_m))^{\perp}$, and let also $v \in S_1$. By definition, $\langle v, v_i \rangle = 0$ for all v_i. Now if $u \in S_2$, that means that $u = \sum_{i=1}^{m} \alpha_i v_i$. Therefore, $\langle v, u \rangle = \langle v, \sum_{i=1}^{m} \alpha_i v_i \rangle = \sum_{i=1}^{m} \overline{\alpha_i} \langle v, v_i \rangle = 0$, and since u was chosen arbitrarily, it holds that $v \in S_2$. Therefore, $S_1 \subset S_2$.
Now, let $v \in S_2$, in which case it is true that $\langle v, \sum_{i=1}^{m} \alpha_i v_i \rangle$ for any choice of α_i. More specifically, if we choose exactly one α_i to be 1 and the rest to be 0, we obtain that $\langle v, v_i \rangle = 0$. Since this is true for all i, we conclude that $v \in S_1$, thus that $S_2 \subset S_1$, completing the proof that $S_1 = S_2$.

Exercise 2
Suppose U is a finite-dimensional subspace of V. Prove that $U^{\perp} = \{0\}$ if and only if $U = V$.

Solution.
$\Longleftarrow$: If $U = V$, then we already know that $V^{\perp} = \{0\}$, thus $U^{\perp} = \{0\}$ (the idea is that for a $v \in V^{\perp}$, v needs to be orthogonal to itself, which happens iff $v = 0$).
$\Longrightarrow$: If $U^{\perp} = \{0\}$, let $v \in V$. Observe that because U is finite-dimensional, we know that $V = U \oplus U^{\perp}$. Therefore, $v = u + w$ for a unique choice of $u \in U, w \in U^{\perp}$. Because $U^{\perp} = \{0\}, w = 0$. This means that $v = u \in U$, which means that $V \subset U$, and since $U \subset V$ (subspace), we have that $U = V$.

Exercise 5
Suppose V is finite-dimensional and U is a subspace of V. Show that $P_{U^{\perp}} = I - P_U$, where I is the identity operator on V.

Solution.
Since V is finite-dimensional, every one of its subspaces is also finite-dimensional. This means that $U, U^{\perp}$ are finite-dimensional, and thus it holds that $(U^{\perp})^{\perp} = U$. It is also true that $V = U^{\perp} \oplus U$. This means that any $v \in V$ can be uniquely written as $v = w + u, w \in U^{\perp}, u \in U$. Then, by definition $P_U(v) = u$ and $P_{U^{\perp}}(v) = w = v - u = v - P_U(v) = I(v) - P_U(v)$, which means that indeed $P_{U^{\perp}} = I - P_U$.

Exercise 6
Suppose U, W are finite-dimensional subspaces of V. Prove that $P_U P_W = 0$ if and only if $\langle u, w \rangle = 0$ for all $u \in U, w \in W$.

Solution.
$\Longrightarrow$: If $P_U P_W(v) = 0$ for every $v \in V$, then let us select a random $w \in W$. Because range$(P_W) = W$, there exists some v for which $P_W(v) = w$, which means that it must hold that $P_U(w) = 0$. This in turn means that $w = P_U(w) + x$, where $x \in U^{\perp}$ and $P_U(w) = 0$. Therefore $w = x \in U^{\perp}$, which by definition means that $\langle u, w \rangle = 0$ for every $u \in U$.

$\Longleftarrow$: We need to show that if $\langle u, w\rangle = 0$ for every $u \in U, w \in W$, then $P_U P_W(v) = 0$ for any $v \in V$. Let then v be any vector in V. Then $P_U P_W(v) = P_U(w)$ for some $w \in W$ (by the definition of P_W). The fact that $\langle u, w\rangle = 0$ for all $u \in U$ means that this w is an element of $U^\perp$. Then, the direct sum decomposition $V = U \oplus U^\perp$ yields that $w = 0 + x, x \in U^\perp$, and therefore $P_U(w) = 0$, which means that $P_U P_W(v) = 0$ for an arbitrary $v \in V$, thus $P_U P_W = 0$.

Exercise 7
Suppose V is finite-dimensional and $P \in L(V)$ is such that $P^2 = P$ and every vector in null$P$ is orthogonal to every vector inrange$P$. Prove that there exists a subspace U of V such that $P = P_U$.

Solution.
We begin by observing that because $P^2 = P$, for any $v \in V$ it must hold that $P^2(v) = P(v) \implies P(P(v)) - P(v) = 0 \implies P(P(v) - v) = 0 \implies P(v) - v \in \text{null}P$. By the given hypothesis, it then holds that $\langle P(v) - v, P(w)\rangle = 0$ for any two $v, w \in V$ (since every vector in null$P$ is orthogonal to every vector in P's range).
We need to show that there exists some U subspace of V such that $P_U = P$. This would mean then that $V = U \oplus U^\perp$, and that for every $v \in V$, $v = P_U(v) + w = P(v) + w$, with $w \in U^\perp$ (where we used $P_U = P$). Additionally, this would mean that $\langle P(v), w\rangle = 0$. Observe now that we can write $v = (v - P(v)) + P(v)$, and that we proved above that $\langle P(v) - v, P(v)\rangle = 0$ for any $v \in V$. Furthermore, since P is a function, $P(v)$ is unique and thus there is a unique way of writing v in this manner. If we were then to set $U = \text{range}P$ (which is clearly a subspace), then any $v \in V$ can be written uniquely as $v = P(v) + (v - P(v)$, with these two vectors being orthogonal and $P(v) \in U$, thus it is indeed true that $P(v) = P_U(v)$ for this choice of U.

Exercise 14
Suppose $C_{\mathbf{R}}([-1, 1]$ is the vector space of continuous real-valued functions on the interval [-1, 1] with inner product given by

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx$$

, for $f, g \in C_{\mathbf{R}}([-1, 1]$. Let U be the subspace of $C_{\mathbf{R}}([-1, 1]$ defined by

$$U = \{f \in C_{\mathbf{R}}([-1, 1] : f(0) = 0\}$$

(a) Show that $U^\perp = \{0\}$.
(b) Show that 6.47 and 6.51 do not hold without the finite-dimensional hypothesis.

Solution.
(a) Clearly, $g(x) = 0$ belongs in $U^\perp$ since $\int_{-1}^{1} f(x)0dx = 0$ for every $f \in U$, therefore $\langle f, g\rangle = 0$ for every $f \in U$. We need to show that no other g is in $U^\perp$. Suppose that this were not true, i.e. there exists some $g \in C_{\mathbf{R}}([-1, 1]$ which is not always zero and is such that $\langle f, g\rangle = 0$ for every f in U.
Suppose first that $g(x) \geq 0$ everywhere. Then, consider an f such that $f(0) = 0$ and $f(x) > 0$ everywhere else (e.g. a parabola symmetric around 0). Then, we can clearly see that $\int_{-1}^{1} f(x)g(x)dx > 0$ (because g is not zero everywhere and is continuous). This means that $\langle f, g\rangle > 0$, i.e. $g \notin U^\perp$. A completely symmetric argument (by considering e.g. f as a concave parabola with $f(0) = 0)$) gives us that it cannot be the case that $g(x) \leq 0$ everywhere (with g not zero everywhere).
Therefore, there exist $x_1, x_2 \in [-1, 1]$ such that $g(x_1)g(x_2) < 0$, and since g is continuous, by the intermediate value theorem we have that $g(\xi) = 0$ for some $\xi \in (-1, 1)$. Due to continuity, it must furthermore be the case that $g(x) \geq 0$ in some neighborhood around ξ and that $g(x) \leq 0$ in some neighborhood around ξ, and that g is not always zero in either of these neighborhoods. It is then easy to find a continuous f such that it has the same sign as g in both of these neighborhoods (consider, e.g. a triangular pulse with $f(\xi = 0)$ like we did in 6.B.17), is not zero everywhere in either of these and has $f(0) = 0$ (even if $g(0) \neq 0$, we can still have f have the same sign as g around 0 and have $f(0) = 0$). But then the product $f(x)g(x)$ is always non-negative and is not zero everywhere, while it is furthermore continuous as a product of continuous functions. Then, $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx > 0$, but $f \in U$ and $g \in U^\perp$, which is a contradiction.
Therefore, there cannot exist a non-zero $g \in U^\perp$, thus $U^\perp = \{0\}$.

(b) Since $U\perp = \{0\}$, if 6.47 was true, i.e. if $V = U \oplus U^\perp$, then we would essentially have $V = U$, which is impossible, since, for example, the function $f(x) = 1$ is in $C_{\mathbf{R}}([-1, 1]$ but not in U ($f(0) \neq 0$).
If 6.51 was true, i.e. if $U = (U^\perp)^\perp$, then that would mean that $(\{0\})^\perp = U$. We know, however, that $\{0\}^\perp = V$ (the proof for this never assumed that V was finite-dimensional), thus we would have that $U = V$, which as we showed is impossible.

Chapter 7

Operators on Inner Product Spaces

7.A Self-Adjoint and Normal Operators

Exercise 1
Suppose n is a positive integer. Define $T \in L(\mathbf{F}^n)$ by

$$T(z_1, \ldots, z_n) = (0, z_1, \ldots, z_{n-1})$$

Find a formula for $T^*(z_1, \ldots z_n)$.

Solution.
The simplest way to find a formula for T^* is to work with the matrices that correspond to these transforms. We know that the standard basis of $\mathbf{F}^n$ is orthonormal, so if we can express $M(T)$ with respect to it, we can obtain $M(T^*)$ by taking the conjugate transpose.
Observe that T maps each vector of the form $(0, \ldots, 1, \ldots, 0)$, with 1 in the i-th position and 0 everywhere else, $1 \leq i \leq n-1$ to the vector $(0, \ldots, 0, 1 \ldots, 0)$, with 1 in the $(i+1)$-th position and 0 everywhere. Furthermore, it maps $(0, \ldots, 1)$ to $(0, \ldots, 0)$. We know that the entries of the k-th column of the matrix of a transform wrt. two bases v_i and w_j are the coefficients that express $T(v_k)$ as a linear combination of w_j. In essence, T maps each vector of the standard basis to the next one (if they are ordered in the "natural" way), except for the last one, which it maps to the zero vector. Here, both of these bases are the standard basis of $\mathbf{F}^n$, so it becomes clear that the matrix $M(T)$ is equal to:

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

Therefore, the matrix of T^* will be the conjugate transpose of this matrix. All the entries are real, thus it suffices to take the transpose and obtain that:

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

Then, since $T^*(v) = M(T^*)M(v)$ (again with respect to the standard basis), we obtain the following formula for T^*:

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

> **Exercise 2**
> Suppose $T \in L(V)$ and $\lambda \in \mathbf{F}$. Prove that λ is an eigenvalue of T if and only if $\overline{\lambda}$ is an eigenvalue of T^*.

Solution.
$\implies$: If λ is an eigenvalue of T, then there exists $v \neq 0$ such that $Tv = \lambda v$. Suppose that $\overline{\lambda}$ is not an eigenvalue of T^*. Then, $T^* - \overline{\lambda}I$ is invertible, which means that $\text{null}(T^* - \overline{\lambda}I) = \{0\}$. Observe that $T^* - \overline{\lambda}I = (T - \lambda I)^*$. We also know that $\text{null}S^* = (\text{range}S)^\perp$, for any linear map S.
More specifically, in this case $\{0\} = \text{null}(T^* - \overline{\lambda}I) = \text{null}((T - \lambda I)^*) = (\text{range}(T - \lambda I))^\perp$.
We also know that $\dim(\text{range}S)^\perp = \dim V - \dim(\text{range}S)$ for any linear map S. Then, $\dim(\text{range}(T - \lambda I)) = \dim V - \dim((\text{range}(T - \lambda I)^\perp) = \dim V - 0 = \dim V$. This means that $T - \lambda I$ is surjective, and thus λ is not an eigenvalue of T, which is a contradiction. Therefore, $\overline{\lambda}$ is an eigenvalue of T^*.
$\impliedby$: By exchanging the roles of T and T^*, as well as λ and $\overline{\lambda}$ in the above proof, and using the fact that $\overline{\overline{\lambda}} = \lambda$ and $(T^*)^* = T$, we obtain the other direction of the equivalency, thus completing the proof.

> **Exercise 5**
> Prove that
> $$\text{dim null } T^* = \text{dim null}T + \dim W - \dim V$$
> and
> $$\text{dim range}T^* = \text{dim range}T$$
> for every $T \in L(V, W)$.

Solution.
For the first equality, we have that:

$$\text{dim null}T^* = \dim\ (\text{range}T)^\perp = \dim \text{ W} - \text{dim range}T =$$

$$\dim W - (\dim V - \text{dim null}T) = \dim W + \text{dim null}T - \dim V$$

, where we used the fact that $\text{null}T^* = (\text{range}T)^\perp$, $\dim U^\perp = \dim V - \dim U$ for any U subspace of any f.d. inner product space V and the Fundamental Theorem of Linear Maps in that order.
For the second equality, by the Fundamental Theorem of Linear Maps appied on $T^* \in L(W, V), T \in L(V, W)$ and the equality we just proved, we have that:

$$\text{dim range}T^* = \dim \text{ W} - \text{dim null}T^* =$$

$$\dim \text{ W} - (\text{dim null}T + \dim W - \dim V) = \dim V - \text{dim null}T = \text{dim range}T$$

> **Exercise 7**
> Suppose $S, T \in L(V)$ are self-adjoint. Prove that ST is self-adjoint if and only if $ST = TS$.

Solution.
$\implies$: If ST is self-adjoint, it holds that $(ST)^* = ST$. But, we know that $(ST)^* = T^*S^* = TS$, since S, T are self-adjoint. Therefore, $ST = (ST)^* = TS$.
$\impliedby$: If $ST = TS$, then let's compute the adjoint of ST. We have that $(ST)^* = T^*S^* = TS$, since S, T are self-adjoint. This means that $(ST)^* = TS = ST$, which means that ST is self-adjoint.

Exercise 8
Suppose V is a real inner product space. Show that the set of self-adjoint operators on V is a subspace of $L(V)$.

Solution. Let S be the set of adjoint operators on V. We have that:

- The operator $T = 0$ is clearly self-adjoint, since: $\langle Tv, w\rangle = 0$ for any $v, w \in V$, which means that $0 = \langle v, T^*w\rangle$ for any $v, w \in V$, which more specifically means that every T^*w must be orthogonal to itself, thus that $T^*w = 0$ for all w, thus that $T^* = 0 = T$.

- If $T_1, T_2 \in S$, then $T_1^* = T_1, T_2^* = T_2$ (self-adjoint operators). Then, $(T_1 + T_2)^* = T_1^* + T_2^* = T_1 + T_2$, therefore $T_1 + T_2$ is also a self-adjoint operator, therefore $T_1 + T_2 \in S$

- If $T \in S$ and $\lambda \in \mathbf{F} = \mathbf{R}$, then $(\lambda T)^* = \overline{\lambda} T^* = \lambda T$, where we used the fact that λ is real and T is self-adjoint. Thus, λT is also self-adjoint, i.e. $\lambda T \in S$.

Combining the above three facts, we conclude that S is a subspace of $L(V)$.

Exercise 9
Suppose V is a complex inner product space with $V \neq \{0\}$. Show that the set of self-adjoint operators on V is not a subspace of $L(V)$.

Solution.
Since $V \neq \{0\}$ and since V is (in this chapter) finite-dimensional, V has a basis that is non-empty. This means that we can define at least one non-zero operator on V, namely the identity operator. For this operator, we know that $I^* = I$, i.e. I is self-adjoint. If we call S the set of self-adjoint operators like we did in the exercise above, we have then that $I \in S$, thus $S \neq \{0\}$.
Now, consider the operator $T = (1 + i)I$. If S is a subspace, then it must hold that $T \in S$, i.e. T is self-adjoint. But, $T^* = ((1+i)I)^* = \overline{1 + i} I^* = (1-i)I$. Obviously, this operator does not equal T, because this would imply equality of their values on the vectors of the basis (which is non-empty), which would be of the form $(1 + i)v_i, (1 - i)v_i$, which (since $v_i \neq 0$) would imply that $1 + i = 1 - i$ which is of course a contradiction.
Thus, S is not a subspace of $L(V)$.

Exercise 11
Suppose $P \in L(V)$ is such that $P^2 = P$. Prove that there is a subspace U of V such that $P = P_U$ if and only if P is self-adjoint.

Solution. $\Longrightarrow$: If there exists such a subspace, then for any v in V we can write $v = P_U(v) + x = P(v) + x$, where $x \in U^\perp$. Let us then consider the quantity $\langle P(v), w\rangle - \langle v, P(w)\rangle$ for any two $v, w \in V$. If we can show that this always zero, then this implies that $P = P^*$ due to the definition of the adjoint. We have that $v = P(v) + x_1, w = P(w) + x_2$, where $P(v), P(w) \in U, x_1, x_2 \in U^\perp$. Then:

$$\langle P(v), w\rangle - \langle v, P(w)\rangle = \langle P(v), P(w) + x_2\rangle - \langle P(v) + x_1, P(w)\rangle$$
$$= \langle P(v), P(w)\rangle - \langle P(v), x_2\rangle - \langle P(v), P(w)\rangle - \langle x_1, P(w)\rangle = -\langle P(v), x_2\rangle - \langle x_1, P(w)\rangle$$

Because $x_1, x_2 \in U^\perp, P(w), P(v) \in U$, it is true that $\langle P(v), x_2\rangle = 0, \langle x_1, P(w)\rangle = 0$, thus the entire expression above is zero, thus $\langle P(v), w\rangle = \langle v, P(w)\rangle$ for any two $v, w \in V$, thus P is self-adjoint.
$\Longleftarrow$: If $P = P^*$, i.e. P is self-adjoint, we have the following. First, observe that $P^2 = P \implies P(P(v)) = P(v) \implies P(P(v) - v) = 0$ for any $v \in V$. Also, we can write $v = (v - P(v)) + P(v)$. If it was the case that $\langle v - P(v), P(v)\rangle = 0$, by choosing $U = \text{range}P$ we can see from this equation that this U would satisfy the condition that $P = P_U$. Observe now that:

$$\langle P(v), v - P(v)\rangle = \langle v, P^*(v - P(v)\rangle = \langle v, P(v - P(v))\rangle = \langle v, 0\rangle = 0$$

, where we used the fact that $P = P^*$, the definition of the adjoint and the fact that $P(v - P(v)) = -P(P(v) - v) = 0$. By the observations made above, this completes the proof if we choose $U = \text{range}P$.

Exercise 15
Fix $u, x \in V$. Define $T \in L(V)$ by:
$$T(v) = \langle v, u\rangle x$$
for every $v \in V$.
(a) Suppose $\mathbf{F} = \mathbf{R}$. Prove that T is self-adjoint if and only if u, x is linearly dependent.
(b) Prove that T is normal if and only if x, u is linearly dependent.

Solution.
(a) $\implies$: Suppose that T is self-adjoint. Then, for any two $v, w \in V$ we have that:

$$\langle T(v), w\rangle = \langle v, T(w)\rangle \implies \langle\langle v, u\rangle x, w\rangle = \langle v, \langle w, u\rangle x\rangle \implies \langle v, u\rangle\langle x, w\rangle = \langle w, u\rangle\langle v, x\rangle$$

, where in the last step we used the fact that $\mathbf{F} = \mathbf{R}$ to omit the conjugate sign over $\langle w, u\rangle$.
Now, by taking $v = u, w = x$ we obtain that:

$$\langle u, u\rangle\langle x, x\rangle = \langle x, u\rangle\langle u, x\rangle \implies |\langle x, u\rangle|^2 = \langle x, x\rangle\langle u, u\rangle = ||x||^2||u||^2$$

, where we used the fact that in real vector spaces $\langle x, u\rangle = \langle u, x\rangle$. Observe that this implies that $|\langle x, u\rangle| = ||x|| \cdot ||u||$, i.e. that the Cauchy-Schwarz inequality is in this case an equality. We know that this is true if and only if one of x, u is a scalar multiple of the other, which for a list of two vectors is equivalent to them being linearly dependent.
$\impliedby$: Suppose now that x, u is linearly dependent. Then, $u = \lambda x$ for some $\lambda \in \mathbf{R}$, except possibly if $x = 0$, but in that case $T(v) = 0$ which is clearly self-adjoint (we also showed this in a previous exercise). Therefore we can move forward assuming that $u = \lambda x, \lambda \in \mathbf{R}$. We then have that, for any two $v, w \in V$:

$$\langle Tv, w\rangle = \langle\langle v, u\rangle x, w\rangle = \langle v, u\rangle\langle x, w\rangle = \langle v, \lambda x\rangle\langle x, w\rangle = \lambda\langle v, x\rangle\langle x, w\rangle$$

$$\langle v, Tw\rangle = \langle v, \langle w, u\rangle x\rangle = \langle w, u\rangle\langle v, x\rangle = \langle w, \lambda x\rangle\langle v, x\rangle = \lambda\langle w, x\rangle\langle v, x\rangle = \lambda\langle v, x\rangle\langle x, w\rangle$$

, where we've again made use of the first and second-slot homogeneity of real inner products, as well as their symmetry. The two equations above show that for any $v, w \in V, \langle Tv, w\rangle = \langle v, Tw\rangle$, thus that T is self-adjoint.
(b) $\implies$: Suppose that T is normal. We know that this means that for any $v \in V, T^*T(v) = TT^*(v)$. From this we obtain that:

$$\langle v, TT^*(v)\rangle = \langle v, T^*T(v)\rangle \implies \langle v, \langle T^*(v), u\rangle x\rangle = \langle T(v), T(v)\rangle \implies \langle v, \langle T^*(v), u\rangle x\rangle = \langle\langle v, u\rangle x, \langle v, u\rangle x\rangle$$
$$\implies \overline{\langle T^*(v), u\rangle}\langle v, x\rangle = \langle v, u\rangle\overline{\langle v, u\rangle}\langle x, x\rangle \implies \langle u, T^*(v)\rangle\langle v, x\rangle = \langle v, u\rangle\langle u, v\rangle\langle x, x\rangle$$
$$\implies \langle T(u), v\rangle\langle v, x\rangle = \langle v, u\rangle\langle u, v\rangle\langle x, x\rangle \implies \langle\langle u, u\rangle x, v\rangle\langle v, x\rangle = \langle v, u\rangle\langle u, v\rangle\langle x, x\rangle$$
$$\implies \langle u, u\rangle\langle x, v\rangle\langle v, x\rangle = \langle v, u\rangle\langle u, v\rangle\langle x, x\rangle$$

If we now set $v = u$ in this final equality, we obtain that:

$$||u||^2\langle x, u\rangle\langle u, x\rangle = ||u||^2||u||^2||x||^2 \implies ||u||^2|\langle x, u\rangle|^2 = ||u||^4||x||^2$$

We know that $||u||^2 = 0$ iff $u = 0$, in which case u, x is always linearly dependent. If $u \neq 0$, dividing the above equation by $||u||^2$ gives us again equality for the Cauchy-Schwarz inequality applied to x, u, which is equivalent to one being a scalar multiple of the other, i.e. x, u being linearly dependent.
$\impliedby$: If x, u is linearly dependent, then either $u = \lambda x$ for some $\lambda \in \mathbf{F}$ or $x = 0$ (or both). In the second case, $T = 0$, thus T is self-adjoint and thus normal. In the first case, we assume that $x \neq 0$ and we obtain that:

$$\langle T^*(v), T^*(v)\rangle = \langle v, TT^*(v)\rangle = \langle v, \langle T^*(v), u\rangle x\rangle = \langle u, T^*(v)\rangle\langle v, x\rangle = \langle T(u), v\rangle\langle v, x\rangle = \langle\langle u, u\rangle x, v\rangle\langle v, x\rangle$$

$$= \langle u, u\rangle\langle x, v\rangle\langle v, x\rangle = \langle \lambda x, \lambda x, \rangle\langle x, v\rangle\langle v, x\rangle = |\lambda|^2||x||^2\langle x, v\rangle\langle v, x\rangle$$

$$\langle T(v), T(v)\rangle = \langle\langle v, u\rangle x, \langle v, u\rangle x\rangle = \langle v, u\rangle\langle u, v\rangle\langle x, x\rangle = \langle v, \lambda x\rangle\langle \lambda x, v\rangle||x||^2 = |\lambda|^2\langle v, x\rangle\langle v, x\rangle||x||^2$$

We observe thus that $||T(v)|| = ||T^*(v)||$ for all $v \in V$, which is equivalent to T being normal.

Exercise 16
Suppose $T \in L(V)$ is normal. Prove that

$$\text{range}T = \text{range}T^*$$

Solution.
First, recall that $\text{range}T^* = (\text{null}T)^\perp$ and that $\text{range}T = (\text{null}T^*)^\perp$. If could show that $\text{null}T = \text{null}T^*$, then this would imply equality of their orthogonal complements as well, which in turn would imply what is being asked for by the exercise. Suppose therefore that $v \in \text{null}T$. Then $T(v) = 0$, and because T is normal, we know that:

$$||T(v)|| = ||T^*(v)|| \implies 0 = ||T^*(v)||$$

, which from the properties of inner products we know is equivalent to $T^*(v) = 0$, i.e. $v \in \text{null}T^*$. Therefore $\text{null} \subset \text{null}T^*$. A completely symmetric argument yields that $\text{null}T^* \subset \text{null}T$, thus $\text{null}T = \text{null}T^*$, and as previously explained, this directly implies that $\text{range}T = \text{range}T^*$.

Exercise 18
Prove or give a counterexample: If $T \in L(V)$ and there exists an orthonormal basis $e_1, \dots, e_n$ of V such that $||T(e_j)|| = ||T^*(e_j)||$ for each j, then T is normal.

Solution.
This is not, in general, true. Let's consider the following case: $\mathbf{F} = \mathbf{R}, V = \mathbf{R}^3, e_i$ the standard basis of $\mathbf{R}^3$ and $T \in L(V)$ such that $M(T) = \begin{pmatrix} 1 & 2 & 0 \\ \sqrt{2} & 1 & 2 \\ \sqrt{2} & \sqrt{2} & 0 \end{pmatrix}$ with respect to the standard basis. In this case, observe the following:

$$M(T^*) = \begin{pmatrix} 1 & \sqrt{2} & \sqrt{2} \\ 2 & 1 & \sqrt{2} \\ 0 & 2 & 0 \end{pmatrix}$$

$$Tv_1 = \begin{pmatrix} 1 & 2 & 0 \\ \sqrt{2} & 1 & 2 \\ \sqrt{2} & \sqrt{2} & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ \sqrt{2} \\ \sqrt{2} \end{pmatrix}, Tv_2 = \begin{pmatrix} 1 & 2 & 0 \\ \sqrt{2} & 1 & 2 \\ \sqrt{2} & \sqrt{2} & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ \sqrt{2} \end{pmatrix},$$

$$Tv_3 = \begin{pmatrix} 1 & 2 & 0 \\ \sqrt{2} & 1 & 2 \\ \sqrt{2} & \sqrt{2} & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$$

$$T^*v_1 = \begin{pmatrix} 1 & \sqrt{2} & \sqrt{2} \\ 2 & 1 & \sqrt{2} \\ 0 & 2 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, T^*v_2 = \begin{pmatrix} 1 & \sqrt{2} & \sqrt{2} \\ 2 & 1 & \sqrt{2} \\ 0 & 2 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \sqrt{2} \\ 1 \\ 2 \end{pmatrix},$$

$$T^*v_3 = \begin{pmatrix} 1 & \sqrt{2} & \sqrt{2} \\ 2 & 1 & \sqrt{2} \\ 0 & 2 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2} \\ \sqrt{2} \\ 0 \end{pmatrix}$$

One can easily observe that for each i, $||Tv_i|| = ||T^*v_i||$. However, if we perform the matrix multiplications that yield TT^*, T^*T, we see that:

$$M(TT^*) = \begin{pmatrix} 5 & 2+\sqrt{2} & 3\sqrt{2} \\ 2+\sqrt{2} & 7 & 2+\sqrt{2} \\ 3\sqrt{2} & 2+\sqrt{2} & 4 \end{pmatrix}, M(T^*T) = \begin{pmatrix} 5 & 4+\sqrt{2} & 2\sqrt{2} \\ 4+\sqrt{2} & 7 & 2 \\ 2\sqrt{2} & 2 & 4 \end{pmatrix}$$

Since the matrices of TT^*, T^*T with respect to the same basis differ, the two linear maps cannot be equal, and thus T cannot be normal despite satisfying the conditions given in the exercise.
Note: The reasoning for finding the example was observing that $||Tv||$ cannot be "decomposed" to a linear combination of $||Te_i||$ if the vectors to which T maps e_i are not orthonormal.

7.B The Spectral Theorem

Exercise 4
Suppose $\mathbf{F} = \mathbf{C}$ and $T \in L(V)$. Prove that T is normal if and only if all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and

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

, where $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T.

Solution.
$\Longrightarrow$: If T is normal, from the (complex) Spectral Theorem we know that T has a diagonal matrix with respect to an orthonormal basis of V that consists of eigenvectors e_i of T. By the conditions equivalent to diagonalizability studied in 5.4, we know that this implies that $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$, where $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T.
Now, if v, w are eigenvectors that correspond to distinct eigenvalues, this means that $v \in E(\lambda_i, T), w \in E(\lambda_j, T)$ for some $i, j, i \neq j$. If $e_{i,1}, \ldots, e_{i,k}$ and $e_{j,1}, \ldots, e_{j,l}$ are the basis vectors of the orthonormal basis that are in $E(\lambda_i, T), w \in E(\lambda_j, T)$ respectively, then v, w can be written as linear combinations of those two lists respectively (since they form bases of the two subspaces). Because the basis e_i is orthonormal, each pair $e_{i,x}, e_{i,y}$ is orthogonal, and thus the vectors v, w are also orthogonal.
$\Longleftarrow$: Again, by the conditions equivalent to diagonalizability we know that if $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$, where $\lambda_1, \ldots, \lambda_m$ are the distinct eigenvalues of T, then T has a diagonal matrix with respect to a basis of V consisting of eigenvectors of T.
Call the vectors of this basis $v_1, \ldots v_n$, and group them into sub-lists based on which $E(\lambda_i, T)$ they belong in (they cannot belong in more than one because we know that eigenvectors corresponding to distinct eigenvalues are linearly independent).
For each of those sub-lists, say $v_{i,1}, \ldots v_{i,k} \in E(\lambda_i, T)$, apply the Gram-Schmidt procedure to obtain an orthonormal basis $e_{i,1}, \ldots, e_{i,k}$ of $E(\lambda_i, T)$. Then, concatenate these bases into a new list $e_1, \ldots, e_n$. Because of the Gram-Schmidt procedure, each e_i has norm equal to one. Furthermore, two distinct e_i, e_j are always orthogonal, because either they correspond to the same eigenvalue, in which case the Gram-Schmidt procedure guarantees that they are orthogonal, or they correspond to distinct eigenvalues, in which case the exercise's hypothesis guarantees that they are orthogonal. Therefore this new list is an orthonormal basis of V consisting —by construction— of eigenvectors of T. Thus, by the conditions equivalent to diagonalizability, T has a diagonal matrix with respect to it and by the Complex Spectral Theorem this implies that T is normal.

Exercise 5
Suppose $\mathbf{F} = \mathbf{R}$ and $T \in L(V)$. Prove that T is self-adjoint if and only if all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and

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

, where $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T.

Solution.
$\Longrightarrow$: If T is self-adjoint, from the (real) Spectral Theorem we know that T has a diagonal matrix with respect to an orthonormal basis of V that consists of eigenvectors e_i of T. Therefore, each $v \in V$ can be written uniquely as $v = \sum_{j=1}^{\dim V} \alpha_j e_j$. Observe also that the e_i can be grouped into sub-lists depending on which $E(\lambda_i, T)$ they belong in (they belong in precisely one because eigenvectors corresponding to distinct eigenvalues must be linearly independent).
Suppose that we call $e_{i,1}, \dots e_{i,k_i} \in E(\lambda_i, T)$ the list of all of the basis vectors that are in $E(\lambda_i, T)$. We will prove that each of those sub-lists is a basis of the corresponding eigenspace. Clearly, it is linearly independent as a sub-list of a linearly independent list. Suppose now that it it is not a spanning list for some i. Then $\dim E(\lambda_i, T) > k_i$. Furthermore, for any other j, $\dim E(\lambda_j, T) \geq k_j$ due to linear independence. Thus $\dim E(\lambda_i, T) + \sum_{j \neq i} \dim E(\lambda_j, T) > k_i + \sum_{j \neq i} k_j$. We know, however, that due to the direct sum, $\dim V$ equals the left side of the inequality, whereas the right side equals precisely the length of the basis e_i, which is again $\dim V$, which is a contradiction. Therefore each of the sub-lists is also a spanning list of the corresponding eigenspace, and hence a basis of it.
Therefore, each $v \in V$ can be written as $v = \sum_{i=1}^{\dim V} v_i$ where $v_i \in E(\lambda_i, T)$ is obtained by summing the terms $a_j e_j$ of the previous sum for which $e_j \in E(\lambda_i, T)$. Additionaly, the only way to write 0 as a sum of such v_i is by taking each $v_i = 0$. This is because e_i is a basis, thus $0 = \sum_{j=1}^{\dim V} \alpha_j e_j \iff a_j = 0$ for all j, which implies also that each $v_i = 0$, by their definition. From this we conclude that indeed $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$.
For any two eigenvectors v, w corresponding to distinct eigenvalues, we can prove that they are orthogonal in exactly the same way we did in exercise 4: they belong in different eigenspaces, thus they are written as linear combinations of two disjoint sub-lists of e_i, and since these are all orthogonal, v, w are orthogonal too.
$\Longleftarrow$: If $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$, where $\lambda_1, \dots, \lambda_m$ denote the distinct eigenvalues of T, and eigenvectors corresponding to distinct eigenvalues are orthogonal, then we have the following. For each $E(\lambda_i, T)$, we can find an orthonormal basis $e_{i,1}, \dots, e_{i,k_i}$ (we know how to find a basis, and then we can apply the Gram-Schmidt procedure to it). Now, due to the direct sum given, any vector $v \in V$ can be written uniquely as $v = \sum_{i=1} \dim V v_i, v_i \in E(\lambda_i, T)$, and each v_i can be written uniquely as a linear combination of $e_{i,1}, \dots, e_{i,k_i}$. From this we conclude that the concatenation of these bases of each $E(\lambda_i, T)$ is a basis of V.
Furthermore, any (distinct) two of its vectors are orthogonal because they either belong in the same eigenspace, in which case the Gram-Schmidt procedure guarantees they are orthogonal, or they belong in different eigenspaces, in which case the given hypothesis guarantees orthogonality. Again due to the Gram-Schmidt procedure(s), each of the basis' vectors is normal. Thus there exists an orthonormal basis of V consisting of eigenvectors of T, and by the real Spectral Theorem this is equivalent to T being self-adjoint.

Exercise 6
Prove that a normal operator on a complex inner product space is self-adjoint if and only if all its eigenvalues are real.

Solution.
$\Longrightarrow$: If an operator T on a complex i.p. space V is normal and self-adjoint, we have the following. By the complex Spectral Theorem, there exists an orthonormal basis of V with respect to which $M(T)$ is diagonal. We know that since this basis is orthonormal, $M(T^*)$ (with respect to the same basis) equals the conjugate transpose of $M(T)$. We have, furthermore, that:

$$M(T) = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & \dots & \dots & a_{nn} \end{pmatrix}, M(T^*) = \begin{pmatrix} \overline{a_{11}} & 0 & \dots & 0 \\ 0 & \overline{a_{22}} & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & \dots & \dots & \overline{a_{nn}} \end{pmatrix}$$

Because $M(T) = M(T^*)$, we conclude that $a_{ii} = \overline{a_{ii}}$ for all i, i.e. that a_{ii} is real. We know, however, that for a diagonal $M(T)$, the values on its diagonal equal precisely the eigenvalues of T. From this we conclude that every eigenvalue of T is real.

$\Longleftarrow$: If T is a normal operator on a complex i.p. space such that all of its eigenvalues are real, we have the following. Again by the complex Spectral Theorem, there exists an orthonormal basis of V with respect to which $M(T)$ is diagonal. And again, we know that the values on the diagonal of $M(T)$ are precisely the eigenvalues of T. Thus, $M(T)$ is of the same form as we described in the " $\Longrightarrow$ " direction, with each $a_{ii} \in \mathbf{R}$. Clearly, this implies that $M(T^*)$, which again —with respect to this basis— equals the conjugate transpose of $M(T)$, must be equal to $M(T)$. From this we conclude that T is indeed self-adjoint.

Exercise 9
Suppose V is a complex inner product space. Prove that every normal operator on V has a square root. (An operator $S \in L(V)$ is called a **square root** of $T \in L(V)$ if $S^2 = T$.)

Solution.
Suppose that T is a normal operator on V. By the complex Spectral Theorem, we know that there exists an orthonormal basis of V such that with respect to it, $M(T)$ is diagonal. Therefore, $M(T)$ is of the form:

$$M(T) = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & \dots & \dots & a_{nn} \end{pmatrix}$$

Since V is a complex inner product space, $\sqrt{a_{ii}}$ is well-defined for every i. Consider then, the operator S whose matrix $M(S)$ with respect to this basis equals:

$$M(S) = \begin{pmatrix} \sqrt{a_{11}} & 0 & \dots & 0 \\ 0 & \sqrt{a_{22}} & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & \dots & \dots & \sqrt{a_{nn}} \end{pmatrix}$$

Note that such an operator S is uniquely defined due to $\mathbf{C}^{\dim V, \dim V}$ being isomorphic to $L(V)$. Observe then, that $M(S)M(S) = M(T)$, and again due to the above isomorphic spaces, and the fact that the matrix of $SS = S^2$ is equal to $M(S)M(S)$, this means that $S^2 = T$. Thus, S is indeed a square root of T.

Exercise 11
Prove or give a counterexample: every self-adjoint operator on V has a cube root. (An operator $S \in L(V)$ is called a **cube root** of $T \in L(V)$ if $S^3 = T$.)

Solution.
If the vector space is real, from the real spectral theorem we know that the self-adjoint T has a diagonal matrix with respect to some orthonormal basis of V. If the vector space if complex, we proved in exercise (6) that a self-adjoint operator has a diagonal matrix whose elements of the diagonal (eigenvalues) are all real, again with respect to some orthonormal basis. In any case, we can therefore write that:

$$M(T) = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & \dots & \dots & a_{nn} \end{pmatrix}$$

, with respect to some orthonormal basis of V and with each $a_{ii} \in \mathbf{R}$. Due to this, we can define $b_{ii} = (a_{ii})^{\frac{1}{3}}$ for each i (since 3 is odd and $a_{ii} \in \mathbf{R}$, the equation $x^3 = a_{ii}$ always has a solution in $\mathbf{R}$). Observe, then, that if we define:

$$M(S) = \begin{pmatrix} b_{11} & 0 & \dots & 0 \\ 0 & b_{22} & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & \dots & \dots & b_{nn} \end{pmatrix}$$

, then $M(S)M(S)M(S) = M(T)$, and since M is an isomorphism between $L(V)$ and $\mathbf{F}^{\dim V, \dim V}$, there exists a corresponding $S \in L(V)$ such that $S^3 = SSS = T$.

7.C Positive Operators and Isometries

Exercise 1
Prove or give a counterexample: If $T \in L(V)$ is self-adjoint and there exists an orthonormal basis $e_1, \dots, e_n$ of V such that $\langle Te_j, e_j \rangle \geq 0$ for each j, then T is a positive operator.

Solution.
This is not, in general, true. Let us present a counterexample. Let T be the operator in $L(\mathbf{C}^2)$ such that $Te_1 = e_2, Te_2 = e_1$ for the standard basis e_1, e_2 of $\mathbf{C}^2$. With respect to this basis, we have that:

$$M(T) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Observe that by taking the complex conjugate of $M(T)$ we obtain $M(T^*) = M(T)$ (we can do this because the basis is orthonormal), from which we conclude that T is self-adjoint. Furthermore, by the definition of T we have that $\langle Te_i, e_i \rangle = 0 \geq 0$ for all e_i. Observe, however, that for example it is true that $T\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \implies \langle T\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rangle = -2 < 0$, therefore T is not a positive operator.

Exercise 4
Suppose $T \in L(V, W)$. Prove that T^*T is a positive operator on V and TT^* is a positive operator on W.

Solution.
Since $T \in L(V, W), T^* \in L(W, V)$, we have that $T^*T \in L(V), TT^* \in L(W)$. Furthermore, for any $v \in V$ it holds that:

$$\langle T^*Tv, v \rangle = \langle Tv, (T^*)^*v \rangle = \langle Tv, Tv \rangle = ||Tv||^2 \geq 0$$

, where we used the definition of the adjoint and the fact that $(T^*)^* = T$. Therefore, we've shown that T^*T is a positive operator on V. In the same manner, for any $w \in W$:

$$\langle TT^*w, w \rangle = \langle T^*w, T^*w \rangle = ||T^*w||^2 \geq 0$$

, and therefore we also proved that TT^* is an operator on W.

Exercise 5
Prove that the sum of two positive operators on V is positive.

Solution.
Let T_1, T_2 be two positive operators on V. Since they are positive, they are also self-adjoint, i.e. $T_1^* = T_1, T_2^* = T_2$. Let then $T = T_1 + T_2$. We know that $T^* = (T_1 + T_2)^* = T_1^* + T_2^* = T_1 + T_2$, i.e. T is also self-adjoint. Furthermore, for any $v \in V$ we have the following:

$$\langle Tv, v \rangle = \langle (T_1 + T_2)(v), v \rangle = \langle T_1v + T_2v, v \rangle = \langle T_1v, v \rangle + \langle T_2v, v \rangle$$

Because T_1, T_2 are positive, we have that $\langle T_1v, v \rangle \geq 0, \langle T_2v, v \rangle \geq 0$. Therefore, we conclude that $\langle Tv, v \rangle \geq 0$ (sum of nonnegatives), which, coupled with the fact that T is self-adjoint, yields that T is also positive. Therefore, the sum of two positive operators is positive.

Exercise 6
Suppose T is positive. Prove that T^k is positive for every positive integer k.

Solution.
T is positive, therefore T is also self-adjoint. Let's first prove that if this is true, then T^k is also self-adjoint for any positive integer k. We can do this by induction on k. For $k = 1$, it clearly holds. Suppose now that it holds for $k \geq 1$, i.e. $(T^k)^* = T^k$. Then we have that, for any $v \in V$:

$$\langle T^{k+1}v, v\rangle = \langle T(T^k(v)), v\rangle = \langle T^k(v), T^*v\rangle = \langle v, (T^k)^*(Tv)\rangle = \langle v, T^k(T(v))\rangle = \langle v, T^{k+1}(v)\rangle$$

, where we used the self-adjointness of T, T^k, and we thus proved that T^{k+1} is also self-adjoint.
By the (real or complex, depending on the associated field) Spectral Theorem, we have that there exists a basis of V consisting of orthonormal $v_1, \ldots, v_n$ eigenvectors of T, and furthermore, becuase T is positive, all of the associated eigenvalues $\lambda_1, \ldots, \lambda_n$ are positive. Furthermore, for any v_i it holds that: $T^k(v_i) = T(T(\ldots(T(v)\ldots) = (\lambda_i)^k v_i$, therefore v_i is an eigenvector of T^k and the corresponding eigenvalue is $\lambda_i^k \geq 0$. T^k can have at most n distinct eigenvalues (since $\dim V = n$ from the assumption that T is self-adjoint and since eigenvectors corresponding to distinct eigenvalues are linearly independent), therefore every eigenvalue of T^k is positive and therefore T^k is positive.

Exercise 7
Suppose T is a positive operator on V. Prove that T is invertible if and only if

$$\langle Tv, v\rangle > 0$$

for every $v \in V$ with $v \neq 0$.

Solution.
$\implies$: If T is invertible, then $Tv = 0$ iff $v = 0$. Because T is positive, T is self-adjoint and there exists an orthonormal basis of V consisting of eigenvectors $e_1, \ldots e_n$ of T (Spectral Theorem), with the corresponding eigenvalues being $\lambda_1, \ldots, \lambda_n$ and each eigenvalue $\lambda_i \geq 0$. Observe also that the fact that $Tv = 0 \iff v = 0$ implies that 0 is not an eigenvalue, thus each λ_i is strictly positive. Consider now a non-zero $v \in V$. v can be written as $v = \sum_{i=1}^n a_i e_i$, with at least one $a_i \neq 0$. Then, we have that:

$$\langle Tv, v\rangle = \langle T(\sum_{i=1}^n a_i e_i), \sum_{j=1} n a_j e_j\rangle = \langle \sum_{i=1}^n a_i \lambda_i e_i, \sum_{j=1}^n a_j e_j\rangle = \sum_{i=1}^n a_i \lambda_i \langle e_i, \sum_{j=1}^n a_j e_j\rangle = \sum_{i=1}^n a_i \lambda_i \sum_{j=1}^n \overline{a_j} \langle e_i, e_j\rangle$$

$$= \sum_{i=1}^n a_i \lambda_i \overline{a_i} = \sum_{i=1}^n |a_i|^2 \lambda_i$$

Because every λ_i is strictly positive and at least one a_i is non-zero, we conclude that the above expression is also strictly positive, i.e. $\langle Tv, v\rangle > 0$ for every $v \in V, v \neq 0$.
$\impliedby$: If $\langle Tv, v\rangle > 0$ for every $v \in V, v \neq 0$, let's assume that T is not invertible. Then, there exists a non-zero $x \in V$ such that $Tx = 0$. But then it would be the case that $\langle Tx, x\rangle = 0$ with $x \neq 0$, which contradicts our hypothesis. Therefore, T is invertible.

Exercise 8
Suppose $T \in L(V)$. For $u, v \in V$, define $\langle u, v\rangle_T$ by

$$\langle u, v\rangle_T = \langle Tu, v\rangle$$

Prove that $\langle \cdot, \cdot\rangle_T$ is an inner product on V if and only if T is an invertible positive operator (with respect to the original inner product $\langle \cdot, \cdot\rangle$)

Solution.
$\implies$: If $\langle \cdot, \cdot\rangle_T$ is an inner product on V, it must hold that $\langle u, u\rangle_T \geq 0$ for every $u \in V$ and also that $\langle u, u\rangle_T = 0$ iff $u = 0$. We have, then, that:

$$\langle u, u\rangle_T = \langle Tu, u\rangle_T \geq 0$$

for every $u \in U$. This is one of the conditions for positivity of T, the other being that T is self-adjoint. Observe also that because $\langle \cdot, \cdot \rangle_T$ is an inner product, it must hold that:

$$\langle v, u \rangle_T = \overline{\langle u, v \rangle_T} \implies \langle Tv, u \rangle = \overline{\langle Tu, v \rangle} \implies \langle Tv, u \rangle = \langle v, Tu \rangle$$

, where we used the conjugate symmetry property of both our inner products, and consequently proved that T is self-adjoint, thus completing the proof that T is positive.
Additionally, the fact that $\langle u, u \rangle_T = \langle Tu, u \rangle = 0$ iff $u = 0$ and that $\langle u, u \rangle_T = \langle Tu, u \rangle \geq 0$ implies that whenever $u \neq 0$, $\langle Tu, u \rangle > 0$, which by the previous exercise (and because T is positive) implies that T is invertible.
$\Longleftarrow$: If T is an invertible positive operator, we need to show that $\langle \cdot, \cdot \rangle_T$ is an inner product on V by showing that it has each of the properties that define an inner product. We have that:

- **Additivity in the first slot**: For $u_1, u_2, v \in V$, we have that $\langle u_1 + u_2, v \rangle_T = \langle T(u_1 + u_2), v \rangle = \langle Tu_1 + Tu_2, v \rangle = \langle Tu_1, v \rangle + \langle Tu_2, v \rangle = \langle u_1, v \rangle_T + \langle u_2, v \rangle_T$, where we used the linearity of T and the first-slot additivity of $\langle \cdot, \cdot, \rangle$.

- **Homogeneity in the first slot**: For $u, v \in V$, $\lambda \in \mathbf{F}$, we have that $\langle \lambda u, v \rangle_T = \langle T(\lambda u), v \rangle = \langle \lambda Tu, v \rangle = \lambda \langle Tu, v \rangle = \lambda \langle u, v \rangle_T$, where we used the linearity of T and the first-slot homogeneity of $\langle \cdot, \cdot \rangle$.

- **Positivity**: For $u \in U$, we have that $\langle u, u \rangle_T = \langle Tu, u \rangle \geq 0$, since T is a positive operator.

- **Definiteness**: For $u \in U$, we have that $\langle u, u \rangle_T = 0 \iff \langle Tu, u \rangle = 0$, which, given that T is positive and invertible implies that $u = 0$. This comes from the previous exercise: if we assume u to be non-zero, this exercise guarantees that $\langle Tu, u \rangle > 0$, which would imply that $\langle u, u \rangle_T > 0$. Since it is obviously true that for $u = 0$, $\langle Tu, u \rangle = 0 = \langle u, u \rangle_T$, we conclude that $\langle u, u \rangle_T = 0 \iff u = 0$, i.e. the condition for definiteness.

- **Conjugate symmetry**: For $u, v \in V$. we have that $\langle v, u \rangle_T = \langle Tv, u \rangle = \langle v, Tu \rangle = \overline{\langle Tu, v \rangle} = \overline{\langle u, v \rangle_T}$, where we used the fact that T is self-adjoint (since it is positive) and the conjugate symmetry of $\langle \cdot, \cdot, \rangle$.

Therefore, $\langle \cdot, \cdot \rangle_T$ is indeed an inner product on V.

7.D Polar Decomposition and Singular Value Decomposition

Exercise 1
Fix $u, x \in V$ with $u \neq 0$. Define $T \in L(V)$ by

$$Tv = \langle v, u \rangle x$$

for every $v \in V$. Prove that

$$\sqrt{T^* T} v = \frac{||x||}{||u||} \langle v, u \rangle u$$

for every $v \in V$.

Solution.
We have that, for every $v, w \in V, \langle Tv, w \rangle = \langle v, T^* w \rangle \implies \langle \langle v, u \rangle x, w \rangle = \langle v, T^* w \rangle \implies \langle v, u \rangle \langle x, w \rangle = \langle v, T^* w \rangle \implies \langle v, \langle w, x \rangle u \rangle - \langle v, T^* w \rangle = 0 \implies \langle v, \langle w, x \rangle u - T^* w \rangle = 0$
Now, for a fixed v this must be true for every $w \in V$ (by the definition of the adjoint). From this we conclude that $T^* w = \langle w, x \rangle u$ (the only vector that is in $V^\perp$ is the zero vector). This now means that:

$$T^* T(v) = T^*(\langle v, u \rangle x) = \langle \langle v, u \rangle x, x \rangle u = \langle v, u \rangle \langle x, x \rangle u = ||x||^2 \langle v, u \rangle u$$

Observe, now, that if we let $Sv = \frac{||x||}{||u||} \langle v, u \rangle u$, then $S(S(v)) = S(\frac{||x||}{||u||} \langle v, u \rangle u) = \frac{||x||}{||u||} \langle \frac{||x||}{||u||} \langle v, u \rangle u, u \rangle u = \frac{||x||^2}{||u||^2} \langle v, u \rangle ||u||^2 u = ||x||^2 \langle v, u \rangle u$, which means that S is a square root of $T^* T$. Furthermore, for any $v \in V$:

$$\langle Sv, v\rangle = \langle \frac{||x||}{||u||}\langle v, u\rangle u, v\rangle = \frac{||x||}{||u||}|\langle v, u\rangle|^2 \geq 0$$

If we can show that S is self-adjoint, then S is positive, and by the uniqueness of positive square roots of positive operators, we'll have proved that $\sqrt{T^*T} = S$. We have that, for any $v, w \in V$:

$$\langle Sv, w\rangle = \langle \frac{||x||}{||u||}\langle v, u\rangle u, w\rangle = \frac{||x||}{||u||}\langle\langle v, u\rangle u, w\rangle = \frac{||x||}{||u||}\langle v, u\rangle\langle u, w\rangle$$

$$\langle v, Sw\rangle = \langle v, \frac{||x||}{||u||}\langle w, u\rangle u\rangle = \frac{||x||}{||u||}\langle v, \langle w, u\rangle u\rangle = \frac{||x||}{||u||}\langle u, w\rangle\langle v, u\rangle$$

, which shows that $\langle Sv, w\rangle = \langle v, Sw\rangle$, therefore S is indeed self-adjoint, thus the proof is complete.

Exercise 3
Suppose $T \in L(V)$. Prove that there exists an isometry $S \in L(V)$ such that

$$T = \sqrt{TT^*}S$$

Solution.
We know that for $T \in L(V)$, $T^* \in L(V)$, and therefore we can apply the Polar Decomposition Theorem to T^*, which means that there exists an isometry $S_a \in L(V)$ such that $T^* = S_a\sqrt{(T^*)^*T^*} = S_a\sqrt{TT^*}$. Then, observe that:

$$T = (T^*)^* = (S_a\sqrt{TT^*})^* = (\sqrt{TT^*})^*S_a^* = \sqrt{TT^*}S_a^{-1}$$

, where we used the fact that $\sqrt{TT^*}$ is self-adjoint (as a positive operator) and that $S_a^* = S_a^{-1}$ since S_a is an isometry. Therefore, if we set $S = S_a^{-1}$, we know that S is also an isometry, and does indeed satisfy the equation $T = \sqrt{TT^*}S$

Exercise 8
Suppose $T \in L(V), S \in L(V)$, S is an isometry and $R \in L(V)$ is a positive operator such that $T = SR$. Prove that $R = \sqrt{T^*T}$
[*The exercise above shows that if we write T as the product of an isometry and a positive operator (as in the Polar Decomposition), then the positive operator equals $\sqrt{T^*T}$.*]

Solution.
By the Polar Decomposition, there exists an isometry S_1 such that $T = S_1\sqrt{T^*T}$. Furthermore, $T = SR$. Therefore, for any $v \in V$:

$$SR(v) = S_1\sqrt{T^*T}(v) \implies ||SR(v)|| = ||S_1\sqrt{T^*T}(v)|| \implies ||R(v)|| = ||\sqrt{T^*T}(v)||$$

, where we used the fact that S, S_1 are isometries. Continuing:

$$||R(v)|| = ||\sqrt{T^*T}(v)|| \implies \langle R(v), R(v)\rangle = \langle\sqrt{T^*T}(v), \sqrt{T^*T}(v)\rangle \implies$$
$$\langle v, RR(v)\rangle = \langle v, \sqrt{T^*T}\sqrt{T^*T}v\rangle \implies \langle v, R^2(v)\rangle = \langle v, T^*T(v)\rangle \implies \langle v, R^2(v) - T^*T(v)\rangle = 0$$

, where we used the definition of the square root of an operator and the properties of self-adjoint operators ($R, \sqrt{T^*T}$ are both self-adjoint as positive operators). Now, since this is true for any v, and since $R^2 - T^*T$ is self-adjoint (this can be shown very easily by taking the adjoint and recalling that R is self-adjoint) this means $R^2 - T^*T$ is the zero operator, therefore that $R^2(v) = T^*T(v)$. This means that R is a square root of T^*T, and since it is also positive, by the uniqueness of positive square roots we have that $R = \sqrt{T^*T}$.

Exercise 9
Suppose $T \in L(V)$. Prove that T is invertible if and only if there exists a unique isometry $S \in L(V)$ such that $T = S\sqrt{T^*T}$.

Solution.
$\Longrightarrow$: Assume that T is invertible. This is equivalent to dim null$T = 0$, which is equivalent to dim null$T^* =$ dim range$T^\perp =$ dim$V -$ dim range$T =$ dim null$T = 0$, i.e. in finite-dimensional vector spaces T being invertible is equivalent to T^* being invertible. By the Polar Decomposition, we know that we can write $T = S\sqrt{T^*T}$ for some isometry S. In particular, let's suppose that we can write $T = S_1\sqrt{T^*T}$ and $T = S_2\sqrt{T^*T}$ for two isometries S_1, S_2. In order to prove that the polar decomposition is unique, we need to prove that $S_1 = S_2$.
Note that we then have $T^* = \sqrt{T^*T}S_1^*, T^* = \sqrt{T^*T}S_2^*$, which implies that for every $v \in V$, $\sqrt{T^*T}((S_1^* - S_2^*)(v)) = 0$, which means that $(S_1^* - S_2^*)(v) \in \text{null}\sqrt{T^*T}$. Observe now that if $v \in \text{null}\sqrt{T^*T}$, then $\sqrt{T^*T}(v) = 0 \implies T^*T(v) = 0 \implies T(v) \in \text{null}T^*$, which in our case means $T(v) = 0$ (since T, T^* are invertible), which means $v \in \text{null}T$, so $v = 0$.
Therefore, we obtain that $(S_1^* - S_2^*)(v) = 0$ for every $v \in V$, which means that $S_1^* = S_2^*$. Clearly then, $S_1 = (S_1^*)^* = (S_2^*)^* = S_2$, which proves that the isometry is unique.
$\Longleftarrow$: Suppose now that for some $T \in L(V)$, we can write $T = S\sqrt{T^*T}$ for a unique isometry S. Suppose that T is not invertible. Then, dim null$T > 0$, and by the observations that we made above this also means that dim null$T^* > 0$. In the proof of the Polar Decomposition, we observed that $||Tv|| = ||\sqrt{T^*T}v||$ and, additionally, that dim range$\sqrt{T^*T} =$ dim range$T$, dim range$\sqrt{T^*T}^\perp =$ dim range$T^\perp =$ dim null$T^* > 0$.
The important point is that the isometry S is fully defined by its values on the orthonormal bases of range$\sqrt{T^*T}$ and range$\sqrt{T^*T}^\perp$ (since V is a direct sum of those). Define then an isometry S_{11} such that $S_{11}(\sqrt{T^*T}v) = Tv$ for all $v \in \text{range}\sqrt{T^*T}$ and an isometry S_12 such that it maps an orthonormal basis e_i of range$\sqrt{T^*T}^\perp$ to an orthonormal basis f_i of range$T^\perp$ (these have the same length). If we continue the construction as in the Polar Decomposition proof by defining $S_1(v) = S_{11}(u) + S_{12}(w)$, when $v = u + w$ with $u \in \text{range}\sqrt{T^*T}, w \in \text{range}\sqrt{T^*T}^\perp$, we end up with an isometry that satisfies $T = S\sqrt{T^*T}$ (see the proof for more details).
Observe that, if we repeat the procedure above and define $S_{21} = S_{11}$ and S_{22} such that it maps e_i to $-f_i$, we will again end up with an isometry S_2 that satisfies $T = S\sqrt{T^*T}$. However, $S_1 \neq S_2$ due to the fact that they are not equal in range$\sqrt{T^*T}^\perp$. Thus S would not be unique, which is a contradiction, and therefore T is invertible.

Exercise 10
Suppose $T \in L(V)$ is self-adjoint. Prove that the singular values of T equal the absolute values of the eigenvalues of T, repeated appropriately.

Solution.
T is self-adjoint, which means that $T^* = T$. This means that $T^*T = TT = T^2$. We know that the operator $T^*T = T^2$ is positive. Furthermore, by the characterization of self-adjoint operators based on the Spectral Theorem, we have that T has a diagonal matrix with respect to some orthonormal basis of V, consisting of eigenvectors of T, and additionally that all of the eigenvalues of T are real (diagonal elements). Furthermore, $M(T^2) = M(T)M(T)$ (with respect to the same basis), which gives us that the eigenvalues of T^2 are all real numbers of the form $\lambda_i^2, \lambda_i \in \mathbf{R}$.
We know, however, that the singular values of T equal the nonnegative square roots of the eigenvalues of $T^*T = T^2$, repeated appropriately, i.e. , $\sqrt{\lambda_i^2} = |\lambda_i|$, each repeated dim$E(\lambda_i^2, T^*T)$ times.

Exercise 11
Suppose $T \in L(V)$. Prove that T and T^* have the same singular values.

Solution.
By exercise 17, we know that if the singular value decomposition of T is

$$Tv = s_1\langle v, e_1\rangle f_1 + \ldots + s_n\langle v, e_n\rangle f_n$$

then the singular value decomposition of T^* is

$$T^*v = s_1\langle v, f_1\rangle e_1 + \ldots + s_n\langle v, f_n\rangle e_n$$

It therefore holds that:

$$T^*Tv = T^*(s_1\langle v, e_1\rangle f_1 + \ldots + s_n\langle v, e_n\rangle f_n) = s_1\langle v, e_1\rangle T^*(f_1) + \ldots + s_n\langle v, e_n\rangle T^*(f_n)$$

$$\implies T^*Tv = s_1\langle v, e_1\rangle s_1 e_1 + \ldots + s_n\langle v, e_n\rangle s_n e_n = s_1^2\langle v, e_1\rangle e_1 + \ldots + s_n^2\langle v, e_n\rangle e_n$$

, where we used the orthonormality of f_i. Observe that if we compute $T^*T(e_i)$ we obtain that each of those vectors is an eigenvector of T^*T, with corresponding eigenvalue s_i^2.
If we repeat the exact same procedure for TT^*, we obtain:

$$TT^*v = s_1^2\langle v, f_1\rangle f_1 + \ldots + s_n^2\langle v, f_n\rangle f_n$$

, which again by the same reasoning means that TT^* has s_i^2 as its eigenvalues.
Since T^*T, TT^* have the same eigenvalues, their positive square roots have the same eigenvalues too (by taking square roots). Hence, T and T^* have the same singular values (by definition of the singular values of an operator).

Exercise 13
Suppose $T \in L(V)$. Prove that T is invertible if and only if 0 is not a singular value of T.

Solution.
$\implies$: Suppose that T is invertible. Suppose, then, also, that T has 0 as a singular value, which means that 0 is an eigenvalue of $\sqrt{T^*T}$. Then, there exists $v \in V, v \neq 0$ such that $\sqrt{T^*T} = 0v = 0$. Then, by the Polar Decomposition:

$$T(v) = S\sqrt{T^*T}(v) = S(0) = 0$$

, which means that T is not invertible, which is a contradiction. Therefore, 0 is not a singular value of T.
$\impliedby$: Suppose that 0 is not a singular value of T, which also means that 0 is not an eigenvalue of $\sqrt{T^*T}$. Suppose, then, that T is not invertible, which means that there exists $v \in V, v \neq 0$ such that $T(v) = 0$. Then, by the Polar Decomposition:

$$T(v) = S\sqrt{T^*T}(v) \implies 0 = S\sqrt{T^*T}(v) \implies \sqrt{T^*T}(v) \in \text{null}S$$

Since S is an isometry, null$S = \{0\}$ (otherwise it would map a vector with non-zero norm to the zero vector, contradiction). Therefore, $\sqrt{T^*T}(v) = 0$, and $v \neq 0$, which would mean that 0 is an eigenvalue of $\sqrt{T^*T}$, which is a contradiction. Therefore, T is invertible.

Exercise 14
Suppose $T \in L(V)$. Prove that dim range$T$ equals the number of nonzero singular values of T.

Solution.
By the Singular Value Decomposition, we know that there exist orthonormal bases of V, $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ such that for any $v \in V$ it holds that:

$$Tv = s_1\langle v, e_1\rangle f_1 + \ldots + s_n\langle v, e_n\rangle f_n$$

, with s_i being the singular values of T. We observe that this means that range$T$ is spanned by the vectors $s_1 f_1, \ldots, s_n f_n$: for any $w \in$ range$V$, we can write it as a linear combination of these vectors as indicated above. Suppose now that we discard all f_j for which $s_j = 0$, which means that we discard all vectors of the basis f_i which correspond to zero singular values of T, which means that we discard precisely as many vectors as the number of zero singular values of T. Therefore, we are left with as many vectors as the number of non-zero singular values of T.

By doing this, the span of the remaining $s_i f_i$ does not change, i.e. , it still equals range$T$. Furthermore, these remaining vectors are non-zero multiples of vectors that are linearly independent (since the entire set of f_i is a basis of V, every one of its subsets is also linearly independent). One can easily see that this means that they are also linearly independent: if they were not, we could find a_i not all zero such that $0 = \sum_i a_i(s_i f_i) = \sum_i (a_i s_i) f_i$, and since all s_i are non-zero, we would have that f_i are not linearly independent, contradiction.
Therefore, the remaining vectors $s_i f_i$ span range$T$ and are also linearly independent, which means that they are a basis of range$T$, which means that its dimension indeed equals the number of non-zero singular values of T.

Exercise 15
Suppose $S \in L(V)$. Prove that S is an isometry if and only if all the singular values of S equal 1.

Solution.
$\Longrightarrow$: Suppose S is an isometry. Then, we know that $S^*S = I$. Furthermore, the unique positive square root of I is I. This means that $\sqrt{S^*S} = I$. Clearly, then, all of the eigenvalues of $\sqrt{S^*S}$ equal 1, which means, by definition, that all of the singular values of S equal 1 as well.
$\Longleftarrow$: Suppose that all of the singular values of S equal 1. This means that all of the eigenvalues of $\sqrt{S^*S}$ equal 1. Because this is a positive and hence self-adjoint operator, by the Spectral Theorem there exists an orthonormal basis of eigenvectors of V with respect to which$\sqrt{S^*S}$ has a diagonal matrix. Furthermore, because all of its eigenvalues equal 1, we conclude that this matrix equals the identity matrix (technically, with respect to this basis), which means that $\sqrt{S^*S} = I$ (due to $\mathbf{F}^{n,n}$ being isomorphic to $L(V)$). This in turn means that $S^*S = I$, which we know is equivalent to S being an isometry.

Exercise 17
Suppose $T \in L(V)$ has singular value decomposition given by

$$Tv = s_1\langle v, e_1\rangle f_1 + \ldots + s_n\langle v, e_n\rangle f_n$$

for every $v\ inV$, where $s_1, \ldots, s_n$ are the singular values of T and $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ are orthonormal bases of V.
(a) Prove that if $v \in V$, then

$$T^*v = s_1\langle v, f_1\rangle e_1 + \ldots + s_n\langle v, f_n\rangle e_n$$

(b) Prove that if $v\ inV$ then

$$T^*Tv = s_1^2\langle v, e_1\rangle e_1 + \ldots + s_n^2\langle v, e_n\rangle e_n$$

(c) Prove that if $v \in V$, then

$$\sqrt{T^*T}v = s_1\langle v, e_1\rangle e_1 + \ldots + s_n\langle v, e_n\rangle e_n$$

(d) Suppose T is invertible. Prove that if $v \in V$, then

$$T^{-1}v = \frac{\langle v, f_1\rangle e_1}{s_1} + \ldots + \frac{\langle v, f_n\rangle e_n}{s_n}$$

Solution.
(a) According to the Singular Value Decomposition, the bases $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ are defined such that $\sqrt{T^*T}e_j = s_j e_j$ for every j, and such that $S(e_j) = f_j$ for every j, with $T = S\sqrt{T^*T}$
Furthermore, by taking the adjoint of T^* we have that $T^* = \sqrt{T^*T}^* S^* = \sqrt{T^*T}S^{-1}$. Observe that the definition $Se_i = f_i$ implies that $S^{-1}f_i = e_i$. Furthermore, for any $v \in V$ we have that:

$$v = \langle v, f_1\rangle f_1 + \ldots + \langle v, f_n\rangle f_n$$

By applying S^{-1} and then $\sqrt{T^*T}$ to both sides of the equation, based on their definitions above, we obtain

$$S^{-1}v = \langle v, f_1\rangle e_1 + \ldots + \langle v, f_n\rangle e_n \implies \sqrt{T^*T}S^{-1}v = s_1\langle v, f_1\rangle e_1 + \ldots + s_n\langle v, f_n\rangle e_n$$

, and since $T^* = \sqrt{T^*T}S^{-1}$, the proof is complete.
(b) Observe that, since $\sqrt{T^*T}e_i = s_ie_i$ for every i, we have that $T^*Te_i = \sqrt{T^*T}\sqrt{T^*T}e_i = s_i^2e_i$ for every i. Thus, for every $v \in V$:

$$v = \langle v, e_1\rangle e_1 + \ldots + \langle v, e_n\rangle e_n \implies T^*Tv = s_1^2\langle v, e_1\rangle e_1 + \ldots + s_n^2\langle v, e_n\rangle e_n$$

(c) This is obvious by the definitions of s_i, e_i (eigenvalues of $\sqrt{T^*T}$, and e_i being an orthonormal basis of the corresponding eigenvectors).
(d) Firstly, from exercise 13 we have that if T is invertible, all of its singular values are non-zero. Thus $\frac{1}{s_i}$ is always well-defined. Furthermore, this implies that $\sqrt{T^*T}$ is invertible, while S is invertible due to being an isometry. From this we have that:

$$T^{-1} = (S\sqrt{T^*T})^{-1} = \sqrt{T^*T}^{-1}S^{-1}$$

It is also true that since s_i are all non-zero and e_i are a basis of V, s_ie_i are also a basis of V. Furthermore, due to the invertibility of $\sqrt{T^*T}$ it must hold that $\sqrt{T^*T}^{-1}(s_ie_i) = e_i \implies \sqrt{T^*T}e_i = \frac{e_i}{s_i}$ for every i. Therefore, for any $v \in V$ we have that:

$$v = \langle v, f_1\rangle f_1 + \ldots + \langle v, f_n\rangle f_n \implies S^{-1}v = \langle v, f_1\rangle e_1 + \ldots + \langle v, f_n\rangle e_n$$
$$\implies \sqrt{T^*T}^{-1}S^{-1} = \frac{\langle v, f_1\rangle e_1}{s_1} + \ldots + \frac{\langle v, f_n\rangle e_n}{s_n}$$

, which, combined with the fact that $T^{-1} = \sqrt{T^*T}^{-1}S^{-1}$ completes the proof.

Exercise 18
Suppose $T \in L(V)$. Let $\hat{s}$ denote the smallest singular value of T and let s denote the largest singular value of T.
(a) Prove that $\hat{s}||v|| \leq ||Tv|| \leq s||v||$ for every $v \in V$.
(b) Suppose λ is an eigenvalue of T. Prove that $\hat{s} \leq |\lambda| \leq s$.

Solution.
By the Polar Decomposition, there exists an isometry S such that $T = S\sqrt{T^*T}$. This means that for every $v \in V, ||Tv|| = ||\sqrt{T^*T}v||$. Furthermore, by the definition of singular values, it holds that there exists an orthonormal basis $e_1, \ldots, e_n$ of V such that $\sqrt{T^*T}e_i = s_ie_i$, with s_i being the singular values of T. This means that, for every $v \in V$:

$$||Tv||^2 = ||\sqrt{T^*T}||^2 = ||s_1\langle v, e_1\rangle e_1 + \ldots + s_n\langle v, e_n\rangle e_n||^2 = |s_1|^2|\langle v, e_1\rangle|^2 + \ldots + |s_n|^2|\langle v, e_n\rangle|^2$$
$$\implies ||Tv||^2 \leq s^2|\langle v, e_1\rangle|^2 + \ldots + s^2|\langle v, e_n\rangle|^2 = s^2(\sum_i|\langle v, e_i\rangle|^2) = s^2||v||^2$$

, where we used the fact that, by definition, $s_i \leq s$ for each s_i and the orthonormality of e_i. By taking square roots we obtain $||Tv|| \leq s||v||$.
Furthermore, given that $s_i \geq \hat{s}$, we can also obtain that:

$$||Tv||^2 = |s_1|^2|\langle v, e_1\rangle|^2 + \ldots + |s_n|^2|\langle v, e_n\rangle|^2 \geq \hat{s}^2|\langle v, e_1\rangle|^2 + \ldots + \hat{s}^2|\langle v, e_n\rangle|^2 \geq \hat{s}^2(\sum_i|\langle v, e_i\rangle|^2) = \hat{s}^2||v||^2$$

Again, by taking square roots we obtain that $||Tv|| \geq \hat{s}||v||$.
(b) If λ is an eigenvalue of T, then there exists a non-zero v such that $Tv = \lambda v$. Now, using part (a) and the fact that v has a non-zero norm, we obtain that for this v:

$$\hat{s}||v|| \leq ||Tv|| \leq s||v|| \implies \hat{s}||v|| \leq ||\lambda v|| \leq s||v|| \implies \hat{s} \leq |\lambda| \leq s$$

Exercise 20
Suppose $S, T \in L(V)$. Let s denote the largest singular value of S, let t denote the largest singular value of T and let r denote the largest singular value of $S+T$. Prove that $r \leq s+t$.

Solution.
By the Polar Decomposition, there exist isometries S_1, S_2, S_3 such that $T = S_1\sqrt{T^*T}, S = S_2\sqrt{S^*S}, R = S_3\sqrt{R^*R}$, while it also holds that $R = S+T$. It is therefore true that, for any $v \in V$:

$$||Rv|| = ||\sqrt{R^*R}v|| \implies ||Tv + Sv|| = ||\sqrt{R^*R}v|| \implies ||\sqrt{R^*R}v|| \leq ||Tv|| + ||Sv||$$

, where we used the triangle inequality. Recall from exercise 18 that $||Tv|| \leq t||v||, ||Sv|| \leq s||v||$, due to t, s being the largest singular values of T, S respectively. Furthermore, let now v_r be an eigenvector of $\sqrt{R^*R}$ corresponding to r. Of course, v_r is non-zero and it holds that $||\sqrt{R^*R}v_r|| = r||v_r||$ (r is positive). Using these facts in combination with the inequality above, we obtain that:

$$r||v_r|| \leq t||v_r|| + s||v_r|| \implies r \leq s+t$$

Chapter 8

Operators on Complex Vector Spaces

8.A Generalized Eigenvectors and Nilpotent Operators

Exercise 1
Define $T \in L(\mathbf{C}^2)$ by

$$T(w, z) = (z, 0)$$

Find all generalized eigenvectors of T.

Solution.
With respect to the standard basis of $\mathbf{C}^2$, the matrix of T is:

$$M(T) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

We are interested in the generalized eigenvectors of T. We know that it suffices to examine the non-zero solutions of $(T - \lambda I)^{\dim \mathbf{C}^2} = 0 \implies (T - \lambda I)^2 = 0$ for all $\lambda \in \mathbf{C}$. We have, then, that:

$$M((T - \lambda I)^2) = \left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right) \left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right) = \begin{pmatrix} \lambda^2 & -2\lambda \\ 0 & \lambda^2 \end{pmatrix}$$

Therefore:

$$M(T)M(v) = 0 \implies \begin{pmatrix} \lambda^2 & -2\lambda \\ 0 & \lambda^2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} \lambda^2 v_1 - 2\lambda v_2 \\ \lambda^2 v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

We observe that if $\lambda = 0$, any v satisfies the equation. This means that all $v \in \mathbf{C}^2$ are generalized eigenvectors for $\lambda = 0$, and since generalized eigenvectors that correspond to different eigenvalues must be linearly independent, there can be no other combinations of eigenvalues and generalized eigenvectors.

Exercise 2
Define $T \in L(\mathbf{C}^2$ by

$$T(w, z) = (-z, w)$$

Find the generalized eigenspaces corresponding to the distinct eigenvalues of T.

Solution.
Let us first find the eigenvalues of T. With respect to the standard basis, we have that:

$$M(T) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \implies M(T - \lambda I) = \begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix}$$

Therefore:

$$M(T - \lambda I)M(v) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} -\lambda v_1 - v_2 \\ v_1 - \lambda v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

, which implies that $v_1 = \lambda v_2$ and $v_2 = -\lambda v_1$. Substituting the first of those in the second we obtan that $v_2 = -\lambda^2 v_2 = 0 \implies v_2(1+\lambda^2) = 0$. This gives either $\lambda = i$, or $\lambda = -i$ or $v_2 = 0$. If $v_2 = 0$ then $v_1 = 0$, which is not an eigenvector by definition. Therefore the eigenvalues are precisely $\lambda = i$ and $\lambda = -i$.
$\lambda = i$ means that $v_1 = iv_2$ and $v_2 \in \mathbf{C}$, whereas $\lambda = -i$ means that $v_1 = -iv_2$ and $v_2 \in \mathbf{C}$. This means that $E(i,T) = \text{span}\left(\begin{pmatrix} i \\ 1 \end{pmatrix}\right), E(-i,T) = \text{span}\left(\begin{pmatrix} -i \\ 1 \end{pmatrix}\right)$. Observe now that the number of distinct eigenvalues equals the dimension of the vector space, and that generalized eigenvectors corresponding to distinct eigenvalues are always linearly independent. Therefore, the generalized eigenspaces must equal the corresponding eigenspaces, since they cannot be any "larger" and still contain vectors that linearly independent from vectors of the other generalized eigenspace.

Exercise 3
Suppose $T \in L(V)$ is invertible. Prove that $G(\lambda, T) = G(\frac{1}{\lambda}, T^{-1})$ for every $\lambda \in \mathbf{F}$ with $\lambda \neq 0$.

Solution.
We begin by observing that, for an invertible T and a non-zero λ:

$$(T^{-1} - \frac{1}{\lambda}I)^n = (T^{-1} - \frac{1}{\lambda}T^{-1}T)^n = (T^{-1}(I - \frac{1}{\lambda}T))^n = (-\frac{1}{\lambda}T^{-1}(T - \lambda I))^n =$$

$$(-\lambda^{-1}T^{-1}(T-\lambda I))^n = (\lambda T)^{-1}(T-\lambda I)(\lambda T)^{-1}(T-\lambda I)\dots(\lambda T)^{-1}(T-\lambda I) = (\lambda T)^{-n}(T-\lambda I)^n$$

, where we used the following facts. One, that $(\lambda T)(\frac{1}{\lambda}T^{-1}) = I$, which means that $(\lambda T)^{-1} = \frac{1}{\lambda}T^{-1}$. Two, the definition of negative powers of operators, that is, $T^{-n} = (T^{-1})^n$. Three, the fact that, while the composition of linear maps is not, in general, a commutative operation, here we can in fact exchange the order of the terms, because —due to associativity— we are effectively only composing pairs of operators from the set $\{T, T^{-1}, I\}$, and for those operators composition is indeed commutative.
With this result we can now see that if v is a generalized eigenvector of T for the eigenvalue λ, it holds that $(T - \lambda I)^n(v) = 0$, and therefore also that $(T^{-1} - \frac{1}{\lambda}I)^n(v) = 0$, i.e., v is a generalized eigenvector of T^{-1} for the eigenvalue $\frac{1}{\lambda}$.
If we were to exchange the roles of T and T^{-1} in the result above, we would obtain a symmetric formula for $(T-\lambda I)^n$ as a function of $(T^{-1} - \frac{1}{\lambda}I)^n$, and thus the same argument would yield that any generalized eigenvector of T^{-1} corresponding to $\frac{1}{\lambda}$ is a generalized eigenvector of T corresponding to λ.
We have thus proved that $G(\lambda, T) = G(\frac{1}{\lambda}, T^{-1})$.

Exercise 5
Suppose $T \in L(V)$, m is a positive integer, and $v \in V$ is such that $T^{m-1}v \neq 0$ but $T^m v = 0$. Prove that

$$v, Tv, T^2v, \dots, T^{m-1}v$$

is linearly independent.

Solution.
First, observe that since $T^m v = 0$, we have that for any $k \geq 0, T^{m+k}v = 0$. Second, observe that if $m = 1$ this holds trivially, since $T^{m-1}v = Iv = v$, given by the exercise to be non-zero, and we are asked to prove that v is linearly independent, which of course is then true. Therefore, from now on assume that $m \geq 2$.
To prove that the given list is linearly independent, we must prove that $a_0v + a_1Tv + \dots + a_{m-1}T^{m-1}v = 0$ iff all a_i are zero. Assume that for some values of a_i it holds that this sum is zero. Then, we have that:

$$a_0v + a_1Tv + \dots + a_{m-1}T^{m-1}v = 0 \implies a_0T^{m-1}v + a_1T^m v + \dots + a_{m-1}T^{m+m-2} = T^{m-1}0 = 0$$

, and now observe that every term except the first one is zero, since $m \geq 2$ and $T^{m+k}v = 0$ for all $k \geq 0$. Thus the above equation becomes $a_0T^{m-1}v = 0$, and since $T^{m-1}v \neq 0$, we obtain $a_0 = 0$.
If $m = 2$, this is reduced to $a_1Tv = 0$ which of course yields $a_1 = 0$. Otherwise, by omitting the first term (since $a_0 = 0$) and composing with T^{m-2} we obtain by the exact same procedure (since again every exponent except for the first one is at least m) that $a_1 = 0$. By continuing this process, we eventually obtain that all $a_i = 0$, thus completing the proof that the given list is linearly independent.

Exercise 6
Suppose $T \in L(\mathbf{C}^3)$ is defined by $T(z_1, z_2, z_3) = (z_2, z_3, 0)$. Prove that T has no square root. More precisely, prove that there does not exist $S \in L(\mathbf{C}^3)$ such that $S^2 = T$.

Solution.
Suppose that such an S did exist, thus $S^2 = T$. Then, we know that $\text{null}S \subset \text{null}S^2 = \text{null}T$. We observe now that the nullspace of T has dimension 1: it is precisely the vectors of the form $(x, 0, 0)$. Since $\text{null}S$ is a subspace of $\text{null}T$, and since the nullspace of T has dimension 1, there are only two possibilities. Either $\text{null}S = \{0\}$ or $\text{null}S = \text{null}T$.
Suppose that $\text{null}S = \{0\}$. Then $S(v) \neq 0$ for all $v \neq 0$. But, we have that e.g. $T(1,0,0) = (0,0,0) = S(S(1,0,0))$. This would mean that $S(1,0,0) \in \text{null}S$, thus that $S(1,0,0) = 0, (1,0,0) \neq 0$. This is clearly a contradiction.
Suppose now that $\text{null}S = \text{null}T$. Because $T = S^2$, observe that this also means than $\text{null}S = \text{null}S^n$ for any $n \geq 1$. Furthermore, observe that:

$$T^3(z_1, z_2, z_3) = T^2(z_2, z_3, 0) = T(z_3, 0, 0) = (0, 0, 0)$$

, that is, T^3 is the zero operator, which means $\text{null}T^3 = \mathbf{C}^3$. But $T^3 = S^6$, hence also $\text{null}S = \text{null}S^6 = \mathbf{C}^3$, hence S is the zero operator. This is clearly a contradiction because it would imply that $T = S^2 = 0$.
In both cases, we arrive at a contradiction, thus such an S cannot exist.

Exercise 7
Suppose $N \in L(V)$ is nilpotent. Prove that 0 is the only eigenvalue of N.

Solution.
Because N is nilpotent, we know that there exists a basis of V with respect to which the matrix of N is zero everywhere except possibly above the diagonal. Such a matrix is, of course, an upper triangular matrix. We know, also, that if an operator has an upper triangular matrix with respect to some basis, its eigenvalues are precisely the elements of the diagonal of that matrix. Therefore, 0 is the only eigenvalue of N, precisely because all of the diagonal elements of its matrix are zero.

Exercise 8
Prove or disprove: The set of nilpotent operators on V is a subspace of $L(V)$.

Solution.
Let us examine the case of $V = \mathbf{C}^2$. In particular, consider the operators N_1, N_2 such that $N_1(e_1) = e_2, N_1(e_2) = 0, N_2(e_1) = 0, N_2(e_2) = e_1$, with e_1, e_2 the standard basis of $\mathbf{C}^2$. Clearly, they are both nilpotent: $N_1^2(e_1) = N_1(e_2) = 0, N_1^2(e_2) = N_1(0) = 0, N_2^2(e_1) = N_2(0) = 0, N_2^2(e_2) = N_2(e_1) = 0$. If the set of nilpotent operators was a subspace of $L(V)$, it would be closed under addition. Thus, here $N_1 + N_2$ would also have to be nilpotent.
However, $(N_1+N_2)(e_1) = e_2, (N_1+N_2)(e_2) = e_1$. This means that $(N_1+N_2)^2(e_1) = e_1, (N_1+N_2)^2(e_2) = e_2$, therefore $(N_1 + N_2)^2 = I$, which means that every even power of $N_1 + N_2$ equals the identity operator and every odd power equals $N_1 + N_2$. Neither of those operators is the zero operator, therefore, $N_1 + N_2$ is not nilpotent and the set of nilpotent operators cannot be a subspace.

Exercise 9
Suppose $S, T \in L(V)$ and ST is nilpotent. Prove that TS is nilpotent.

Solution.
Let $n = \dim V$. Because ST is nilpotent, we know that $(ST)^n = 0$. It suffices to show that *some* power of TS is the zero operator. Observe that, for $k \geq 2$, $(TS)^k = T(ST)^{k-1}S$. We can prove this by induction on k:

- For $k = 2$, $(TS)^2 = TSTS = T(ST)S$ by the associativity of linear operators.

- If $(TS)^k = T(ST)^{k-1}S$, we need to show that $(TS)^{k+1} = T(ST)^k S$. Indeed, $(TS)^{k+1} = (TS)^k(TS) = T(ST)^{k-1}S(TS) = T(ST)^{k-1}(ST)S = T(ST)^k S$, where we used the associativity of linear operators and the induction hypothesis.

Having proved this, we have that $(TS)^{n+1}(v) = T(ST)^n S(v) = T(0) = 0S$ for any $v \in V$, since $(ST)^n = 0$. Hence, TS is indeed nilpotent.

Exercise 12
Suppose $N \in L(V)$ and there exists a basis of V with respect to which N has an upper-triangular matrix with only 0's on the diagonal. Prove that N is nilpotent.

Solution.
Let $e_1, e_2, \dots e_n$ be the basis of V with respect to which $M(N)$ is upper-triangular with only zeros on the diagonal. We claim then that $N^k(e_l) = 0$ for all $l \leq k$. If this is true, observe that it would imply that $N^n(e_i) = 0$ for every e_i (n is the length of the basis, and thus the dimension of V), and thus N would be nilpotent, since an operator is fully defined by its values on a basis. We will prove this by "induction" (but only for a finite number of values) on k:

- For $k = 1$, observe that $N(e_1) = 0$, since the first column of $M(N)$ consists of zeros only.
- If this is true for all $k \leq l$, we need to prove that it holds for $k = l + 1$. We have that $N(e_{l+1}) = \sum_{i=1}^{l} M(N)_{i,l+1} e_i$, since from the diagonal element of column $l+1$ and below all entries of $M(N)$ are zero. But then $N^{l+1}(e_{l+1}) = \sum_{i=1}^{l} M(N)_{i,l+1} N^l(e_i)$. Every term of this sum is zero: for $i < l + 1$, by the induction hypothesis $N^l(e_i) = 0$. Thus $N^{l+1}(e_{l+1}) = 0$, while it is also true that $N^{l+1}(e_k) = 0, k < l$, trivially by the induction hypothesis.

This completes the proof of our claim, and, as we observed above, this directly implies that N is nilpotent.

Exercise 13
Suppose V is an inner product space and $N \in L(V)$ is normal and nilpotent. Prove that $N = 0$.

Solution.
Suppose that $N \neq 0$. This means that there exists a $v \in V$ such that $N(v) \neq 0$. Furthermore, because N is nilpotent, it must hold that *some* power of N is zero at v (in fact, we know that $N^{\dim V} = 0$). Let k be the smallest power for which $N^k(v) = 0$. By our definition, $k \geq 2$, because if $k = 1$, then $N(v) = 0$.
We have that:

$$N^k(v) = 0 \implies N(N^{k-1}(v)) = 0 \implies ||N(N^{k-1}(v))|| = 0$$

Because N is normal, $||N(u)|| = ||N^*(u)||$ for any $u \in V$. More specifically:

$$||N(N^{k-1}(v))|| = ||N^*(N^{k-1}(v))|| \implies ||N^*(N^{k-1}(v))|| = 0 \implies N^*(N^{k-1}(v)) = 0$$

This means, therefore, that $N^{k-1}(v) \in \text{null}N^* = (\text{range}N)^\perp$. Observe, however, that $N^{k-1}(v) = N(N^{k-2}(v))$, where N^{k-2} is well-defined since $k \geq 2$. This means that $N^{k-1}(v) \in \text{range}N$, and since $N^{k-1}(v) \in (\text{range}N)^\perp$, $N^{k-1}(v)$ must be orthogonal to itself. This implies that $N^{k-1}(v) = 0$, which, by our definition of k, is a contradiction. Thus, N must be the zero operator.

Exercise 16
Suppose $T \in L(V)$. Show that

$$V = \text{range}T^0 \supset \text{range}T^1 \supset \dots \supset \text{range}T^k \supset \text{range}T^{k+1} \supset \dots$$

Solution.
Because $T^0 = I$, it is clear that $V = \text{range}I = \text{range}T^0$. Now, pick $k \geq 0$. We need to show that $\text{range}T^k \supset \text{range}T^{k+1}$. If $v \in \text{range}T^{k+1}$, there exists $w \in V$ such that $T^{k+1}(w) = v \implies T^k(T(w)) = v$. By definition, this means that $v \in \text{range}T^k$. Hence, $\text{range}T^k \supset \text{range}T^{k+1}$, which completes the proof.

Exercise 17
Suppose $T \in L(V)$ and m is a nonnegative integer such that

$$\text{range}T^m = \text{range}T^{m+1}$$

Prove that $\text{range}T^k = \text{range}T^m$ for all $k > m$.

Solution.
Observe that $\text{range}T^m = \text{range}T^{m+1}$ means that $\dim \text{range}T^m = \dim \text{range}T^{m+1}$. Furthermore, by the Fundamental Theorem of Linear Maps, this yields that $\dim \text{null}T^m = \dim V - \dim \text{range}T^m = \dim V - \dim \text{range}T^{m+1} = \dim \text{null}T^{m+1}$. We also know that $\text{null}T^m \subset \text{null}T^{m+1}$, thus the equality of dimensions implies that, in fact, $\text{null}T^m = \text{null}T^{m+1}$.
We know that this implies that for each $k > m$, $\text{null}T^m \subset \text{null}T^k$. Thus, $\dim \text{null}T^m = \dim \text{null}T^k$, and again from the Fundamental Theorem of Linear Maps, we have that $\dim \text{range}T^m = \dim V - \dim \text{null}T^m = \dim V - \dim \text{null}T^k = \dim \text{range}T^k$. Additionally, in the previous exercise we proved that $\text{range}T^m \supset \text{range}T^k$. Thus, the equality of dimensions implies that $\text{range}T^m = \text{range}T^k$ for all $k > m$.

Exercise 18
Suppose $T \in L(V)$. Let $n = \dim V$. Prove that

$$\text{range}T^n = \text{range}T^{n+1} = \text{range}T^{n+2} = \dots$$

Solution.
We know that for $n = \dim V$, $\text{null}T^n = \text{null}T^{n+1}$. This implies that $\dim \text{null}T^n = \dim \text{null}T^{n+1}$, and by the Fundamental Theorem of Linear Maps we then have that $\dim \text{range}T^n = \dim V - \dim \text{null}T^n = \dim V - \dim \text{null}T^{n+1} = \dim \text{range}T^{n+1}$.
Furthermore, by exercise 16, $\text{range}T^n \supset \text{range}T^{n+1}$. Therefore the equality of dimensions implies that $\text{range}T^n = \text{range}T^{n+1}$. By exercise 17, this yields that $\text{range}T^n = \text{range}T^k$ for all $k > n$.

8.B Decomposition of an Operator

Exercise 1
Suppose V is a complex vector space, $N \in L(V)$ and 0 is the only eigenvalue of N. Prove that N is nilpotent.

Solution.
Since V is a complex vector space, there exists some basis of it with respect to which $M(N)$ is upper triangular. We know then that the values on the diagonal of this matrix must equal the eigenvalues of N. Thus, they must all be zero. By exercise 12 of 8.A, we know that the fact that there exists a basis of V with respect to which N has an upper triangular matrix with only zeros on the diagonal implies that N is nilpotent.

Exercise 3
Suppose $T \in L(V)$. Suppose $S \in L(V)$ is invertible. Prove that T and $S^{-1}TS$ have the same eigenvalues with the same multiplicities.

Solution.
We are interested in the solutions of $(T - \lambda I)^n(v) = 0$ and $(S^{-1}TS - \lambda I)^n(v) = 0$, where $n = \dim V$. More precisely, we want to show that the dimensions of the two nullspaces are equal for any λ. We begin by observing that if λ is an eigenvalue of T, then $Tv = \lambda v$ for a nonzero v. Because S is invertible, it is surjective and injective, thus there exists a unique and nonzero $w \in V$ such that $S(w) = v$. Then, $S^{-1}TS(w) = S^{-1}T(v) = S^{-1}(\lambda v) = \lambda w$, which means that λ is an eigenvalue of $S^{-1}TS$. Similarly, if λ is an eigenvalue of $S^{-1}TS$, there exists nonzero v such that $S^{-1}TS(v) = \lambda v$. This means that $TS(v) = S(\lambda v) = \lambda S(v)$, therefore $S(v)$ (which must be nonzero because S is injective) is an eigenvector of T corresponding to λ. Therefore the two transforms do indeed have the same eigenvalues.

We continue by observing the following:

$$(S^{-1}TS - \lambda I)(v) = S^{-1}(TS - \lambda SI)(v) = (S^{-1}(T - \lambda SIS^{-1})S)(v) = (S^{-1}(T - \lambda I)S)(v)$$

We now claim that $(S^{-1}(T - \lambda I)S)^n = S^{-1}(T - \lambda I)^n S$ for any positive integer n. We prove this by induction on n:

- For $n = 1$, this is clearly true.
- If it holds for $n = k$, we have that $(S^{-1}(T - \lambda I)S)^{k+1} = (S^{-1}(T - \lambda I)S)^k(S^{-1}(T - \lambda I)S) = S^{-1}(T - \lambda I)^k SS^{-1}(T - \lambda I)S = S^{-1}(T - \lambda I)^{k+1}S$

  , where we used the induction hypothesis and the associativity of linear map products.

By the above facts, we have that:

$$(S^{-1}TS - \lambda I)^n(v) = (S^-1(T - \lambda I)^n S)(v)$$

Therefore, if v is a generalized eigenvector of T corresponding to λ, v is non-zero and $(T - \lambda I)^n(v) = 0$, and because S is invertible there exists a non-zero and unique w such that $Sw = v$. Thus:

$$(S^{-1}TS - \lambda I)^n(w) = (S^{-1}(T - \lambda I)^n S)(w) = (S^{-1}(T - \lambda I)^n)(v) = S^{-1}(0) = 0$$

, thus w is a generalized eigenvector of $S^{-1}TS$ corresponding to λ. Now, if w is a generalized eigenvector of $S^{-1}TS$ corresponding to λ, we have that w is non-zero and $(S^{-1}TS - \lambda I)^n(w) = (S^{-1}(T - \lambda I)^n S)(w) = 0$. From this we obtain that:

$$(T - \lambda I)^n S(w) = S(0) = 0$$

Because S is invertible and injective, $S(w)$ is non-zero and of course unique. Thus, it is a generalized eigenvector of T corresponding to λ.
The fact that in both cases we were able to *uniquely* map elements of $G(\lambda, T)$ to elements of $G(\lambda, S^{-1}TS)$ and vice versa leads us to conclude that the two subspaces are isomorphic, and thus have the same dimension. By definition, this means that the eigenvalues of $T, S^{-1}TS$ are not only the same, but that the multiplicities of each of them are also equal.

Exercise 5
Suppose V is a complex vector space and $T \in L(V)$. Prove that V has a basis consisting of eigenvectors of T if and only if every generalized eigenvector of T is an eigenvector of T.

Solution.
$\implies$: If V has a basis consisting of eigenvectors of T, it is also true that $V = E(\lambda_1, T) \oplus \ldots \oplus E(\lambda_m, T)$, where λ_i are the eigenvalues of T. Furthermore, it is also true that $V = G(\lambda_1, T) \oplus \ldots \oplus G(\lambda_m, T)$. Suppose now that $v \in G(\lambda_i, T)$ is a generalized eigenvector of T corresponding to λ_i. From the first direct sum, we obtain that there exist $u_1 \in E(\lambda_1, T), \ldots, u_m \in E(\lambda_m, T)$ such that:

$$v = u_1 + \ldots + u_m \implies 0 = u_1 + \ldots + u_i - v + \ldots + u_m$$

Because each eigenvector of T is obviously a generalized eigenvector of T, $u_k \in G(\lambda_k, T)$. Furthermore, $u_i - v \in G(\lambda_i, T)$, precisely because both of these vectors belong in this subspace. Now, from the second direct sum, it must be true that if $0 = w_1 + \ldots + w_m, w_k \in G(\lambda_k, T)$, then each w_k is zero. From this we obtain that $u_k = 0, k \neq i$ and that $u_i - v = 0 = \implies u_i = v$. This means that v belongs in $E(\lambda_i, T)$, therefore it is indeed an eigenvector of T.
$\impliedby$: It is clear that $E(\lambda_i, T) \subset G(\lambda_i, T)$. Furthermore, if every generalized eigenvector of T corresponding to λ_i is an eigenvector of T, it must be an eigenvector corresponding to λ_i. If it corresponded to some other eigenvalue λ_j, then it would also be a generalized eigenvector for λ_j , which is a contradiction due to generalized eigenvectors corresponding to different eigenvalues being linearly independent. This yields that $G(\lambda_i, T) \subset E(\lambda_i, T)$, therefore $E(\lambda_i, T) = G(\lambda_i, T)$. Because V can always be written as a direct sum of all $G(\lambda_i, T)$, in this case it is also true that it can be written as a direct sum of all $E(\lambda_i, T)$, which we know is equivalent to V having a basis consisting of eigenvectors of T.

Exercise 7
Suppose V is a complex vector space. Prove that every invertible operator on V has a cube root.

Solution.
If we can prove that every operator of the form $I + N$, where N is a nilpotent operator, has a cube root, then we could apply the same reasoning as 8.33 (invertible operators have square roots):
First we'd write V as a direct sum of the generalized eigenspaces of T and observe that none of the eigenvalues equal zero. Furthermore, that the restriction of T in each of those subspaces is equal to a nilpotent operator plus a multiple of the identity. Each of those operators will have a cube root, and this would allow us to define an operator R as the sum of the application of each of the cube roots. This means, of course, that $T = R^3$.
Now, to prove that every operator of the form $I + N$, N a nilpotent operator, has a cube root, we can use the same "trick" as 8.31, inspired instead by the Taylor series of $\sqrt[3]{1+x}$. Because it is true that

$$\sqrt[3]{1+x} = 1 + a_1 x + a_2 x^2 + \dots$$

, and because $N^m = 0$ for some m, we need only consider a finite number of terms in the sum. The rest of the proof is done in the same way as 8.31, the difference being that we cube the expression

$$I + a_1 N + a_2 N^2 + \dots + a_{m-1} N^{m-1}$$

instead of squaring it, and then equate it to $I + N$.

Exercise 10
Suppose $\mathbf{F} = \mathbf{C}$ and $T \in L(V)$. Prove that there exist $D, N \in L(V)$ such that $T = D + N$, the operator D is diagonalizable, N is nilpotent and $DN = ND$.

Solution.
Recall that in a complex vector space, given an operator T with eigenvalues $\lambda_1, \dots, \lambda_m$ with corresponding multiplicities $d_1, \dots, d_m$, there always exists a basis of V with respect to which T has a block diagonal matrix of the form

$$\begin{pmatrix} A_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & A_m \end{pmatrix}$$

, where each A_j is a $d_j \times d_j$ upper triangular matrix of the form

$$A_j = \begin{pmatrix} \lambda_j & \dots & * \\ \vdots & \ddots & 0 \\ 0 & \dots & \lambda_j \end{pmatrix}$$

Let then D, N be the operators whose matrices with respect to this basis equal:

$$M(D) = \operatorname{diag}\{M(T)\}, M(N) = M(T) - M(D)$$

, where we use the notation $\operatorname{diag}\{A\}$ to refer to the diagonal matrix that can be formed by selecting the diagonal of matrix A and setting all other elements to zero. Observe, also, that by these definitions $M(N)$ is an upper triangular matrix with zeros on the diagonal. We know that this means that N is a nilpotent operator. Furthermore, by definition it holds that $T = D + N$. Therefore, we only need to prove that $DN = ND$. To do this, we need only consider the values of DN, ND on the vectors $v_1, \dots, v_n$ comprising the basis we are using.
Due to the way we've selected this basis, we can divide the vectors in groups, depending on which eigenvalue they correspond to as generalized eigenvectors: $v_1, \dots, v_{d_1}$ being the first group, $v_{d_1+1}, \dots, v_{d_2}$ the second, and so on until $v_{d_1+d_2+\dots+d_{m-1}+1}, \dots, v_n$ being the last group. Consider then a vector v_i that belongs in the j-th of these groups, and denote as S_j the indices of the vectors of the group.

Then, due to the structure of D, it holds that $D(v_i) = \lambda_j v_i$. Furthermore, due to the structure of N, it holds that $N(v_i) = \sum_{k\in S_j, k\neq i} c_k v_k$. This is true because, for one, N has the same block structure as $M(T)$, i.e., everything not belonging in one of the diagonal blocks is zero, and also because N has zeros on the diagonal. Additionally, $D(v_k) = \lambda_j v_k$ for $k \in S_j$, because of the structure of the A_j matrices, where all of their diagonal elements are equal to λ_j.
Thus, we finally have that:

$$ND(v_i) = N(\lambda_j v_i) = \lambda_j(\sum_{k\in S_j, k\neq i} c_k v_k)$$

$$DN(v_i) = D(\sum_{k\in S_j, k\neq i} c_k v_k) = \sum_{k\in S_j, k\neq i} c_k D(v_k) = \sum_{k\in S_j, k\neq i} c_k \lambda_j v_k$$

, which means that $ND(v_i) = DN(v_i)$. Because v_i is any vector of the basis, the two operators are indeed equal everywhere, thus completing the proof.

8.C Characteristic and Minimal Polynomials

Exercise 8
Suppose $T \in L(V)$. Prove that T is invertible if and only if the constant term in the minimal polynomial of T is nonzero.

Solution.
To begin, note that we know that T is invertible iff 0 is not an eigenvalue of T. Therefore, we can replace "T is invertible" in the exercise statement with "0 is not an eigenvalue of T" and we'll be proving an equivalent statement.
$\Longrightarrow$: If 0 is not an eigenvalue of T, it is also not a root of the minimal polynomial p of T. Let $p(z) = a_0 + a_1 z + \ldots + a_{m-1} z^{m-1} + z^m$ be the minimal polynomial of T. Then, $p(0) = a_0$. Since 0 is not a root of p, it follows that $a_0 \neq 0$, i.e. the constant term of p is non-zero.
$\Longleftarrow$: If the constant term of the minimal polynomial p of T is non-zero, we observe again that $p(0) = a_0 \neq 0$. Therefore, 0 is not a root of p, and thus cannot be an eigenvalue of T, completing the proof.

Exercise 10
Suppose V is a complex vector space and $T \in L(V)$ is invertible. Let p denote the characteristic polynomial of T and let q denote the characteristic polynomial of T^{-1}. Prove that

$$q(z) = \frac{1}{p(0)} z^{\dim V} p(\frac{1}{z})$$

for all nonzero $z \in \mathbf{C}$.

Solution.
Let $\lambda_1, \ldots, \lambda_m$ be the eigenvalues of T, with corresponding multiplicities $d_1, \ldots, d_m$. We know that because T is invertible, every λ_i is nonzero. Additionally, we know from a previous exercise that T^{-1} has eigenvalues $\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_m}$ with corresponding multiplicities $d_1, \ldots, d_m$. Therefore, the characteristic polynomial of T^{-1} is:

$$q(z) = (z - \frac{1}{\lambda_1})^{d_1} (z - \frac{1}{\lambda_2})^{d_2} \ldots (z - \frac{1}{\lambda_m})^{d_m}$$

By a series of algebraic manipulations, and by assuming $z \neq 0$, we get that:

$$q(z) = (-\frac{1}{\lambda_1}(-\lambda_1 z + 1))^{d_1} (-\frac{1}{\lambda_2}(-\lambda_2 z + 1))^{d_2} \ldots (-\frac{1}{\lambda_m}(-\lambda z + 1))^{d_m} \implies$$

$$q(z) = (-\frac{1}{\lambda_1})^{d_1} (-\frac{1}{\lambda_2})^{d_2} \ldots (-\frac{1}{\lambda_m})^{d_m} (z(\frac{1}{z} - \lambda_1))^{d_1} (z(\frac{1}{z} - \lambda_2))^{d_2} \ldots (z(\frac{1}{z} - \lambda_m))^{d_m} \implies$$

$$q(z) = \frac{1}{(-\lambda_1)^{d_1} (-\lambda_2)^{d_2} \ldots (-\lambda_m)^{d_m}} z^{d_1 + d_2 + \ldots + d_m} (\frac{1}{z} - \lambda_1)^{d_1} (\frac{1}{z} - \lambda_2)^{d_2} \ldots (\frac{1}{z} - \lambda_m)^{d_m}$$

We know that $\sum_i d_i = \dim V$. We also know that if p is the characteristic polynomial of T, then $p(z) = (z-\lambda_1)^{d_1}(z-\lambda_2)^{d_2}\dots(z-\lambda_m)^{d_m}$. We observe then that $p(0) = (-\lambda_1)^{d_1}(-\lambda_2)^{d_2}\dots(-\lambda_m)^{d_m}$ and that for $z \neq 0, p(\frac{1}{z}) = (\frac{1}{z}-\lambda_1)^{d_1}(\frac{1}{z}-\lambda_2)^{d_2}\dots(\frac{1}{z}-\lambda_m)^{d_m}$, leading us to conclude that:

$$q(z) = \frac{1}{p(0)} z^{\dim V} p(\frac{1}{z})$$

Exercise 11
Suppose $T \in L(V)$ is invertible. Prove that there exists a polynomial $p \in P(\mathbf{F})$ such that $T^{-1} = p(T)$.

Solution.
We know that the minimal polynomial p of T is always well-defined. Additionally, in exercise 8 we proved that the constant term of this polynomial is non-zero if and only if T is invertible. Therefore, the minimal polynomial here has the form:

$$p(z) = a_0 + a_1 z + \ldots + a_{m-1} z^{m-1} + z^m, a_0 \neq 0$$

Since $p(T) = 0$, for every $v \in V$ it holds that:

$$a_0 I(v) + a_1 T(v) + \ldots + a_{m-1} T^{m-1}(v) + T^m(v) = 0 \implies a_0 T T^{-1} v + a_1 T(v) + \ldots + a_{m-1} T^{m-1}(v) + T^m(v) = 0$$

$$\implies a_0 T T^{-1}(v) = -a_1 T(v) - \ldots - a_{m-1} T^{m-1}(v) - T^m(v)$$

$$\implies T T^{-1}(v) = -\frac{a_1}{a_0} T(v) - \ldots - \frac{a_{m-1}}{a_0} T^{m-1}(v) - \frac{1}{a_0} T^m(v)$$

$$\implies T^{-1}(v) = -\frac{a_1}{a_0} v - \ldots - \frac{a_{m-1}}{a_0} T^{m-2}(v) - \frac{1}{a_0} T^{m-1}(v)$$

, where we can divide by a_0 since it is not zero. The equation above shows that we've expressed T^{-1} as a polynomial q of T, therefore completing the proof.

Exercise 12
Suppose V is a complex vector space and $T \in L(V)$. Prove that V has a basis consisting of eigenvectors of T if and only if the minimal polynomial of T has no repeated zeros.

Solution.
$\implies$: Suppose that V has a basis consisting of eigenvectors $v_1, \dots, v_n$ of T, each of them corresponding to *some* eigenvalue of T, and let the eigenvalues of T be $\lambda_1, \dots, \lambda_m$. We know that the zeros of the minimal polynomial are precisely λ_i, therefore the minimal polynomial is of the form $p(z) = (z-\lambda_1)^{e_1}(z-\lambda_2)^{e_2}\dots(z-\lambda_m)^{e_m}, e_i > 0$.
Let v be any vector in V. Then, $v \sum_i a_i v_i$. Observe that this means that $(T-\lambda_1 I)(T-\lambda_2 I)\dots(T-\lambda_m I)(\sum_i a_i v_i) = \sum_i a_i (T-\lambda_1 I)\dots(T-\lambda_m I)v_i$. All of the operators in each of the terms of the sum commute. Furthermore, each v_i is an eigenvector of some eigenvalue, and therefore by rearranging these commutative operators so that in each term of the sum the operator $(T-\lambda_j)$, λ_j being the eigenvalue that corresponds to v_i, is applied first, we can see that the sum equals zero. Therefore $(T-\lambda_1 I)(T-\lambda_2 I)\dots(T-\lambda_m I)$ is the zero operator, which leads us to conclude that $p(z) = (z-\lambda_1)\dots(z-\lambda_m)$ must be the minimal polynomial of T (since each exponent should be at least 1).
Clearly, there are no repeated zeros in the minimal polynomial of T.
$\impliedby$: We'll prove this direction by proving the contrapositive, i.e. we will assume that there is at least one repeated zero in the minimal polynomial of T and show that this implies that there exists no basis of V consisting of eigenvectors of T.
Suppose therefore that p, the minimal polynomial of T, is of the form $p(z) = (z-\lambda_1)^{e_1}\dots(z-\lambda_m)^{e_m}$, where at least one exponent is greater than 1, and without loss of generality, suppose that $e_1 > 1$. p can therefore be written as $p(z) = (z-\lambda_1)^{e_1-1}(z-\lambda_1)q(z)$. Because p is the minimal polynomial of T, $(T-\lambda_1 I)q(T)$ cannot be zero everywhere. Therefore, there exists a nonzero v such that $(T-\lambda_1 I)q(T)(v)$ is non-zero.

Additionally, it must hold that $p(T)(v) = 0$. If $q(T)(v) = w$, then w is non-zero and $(T - \lambda_1 I)(w)$ is also non-zero. However, $p(T)(w) = (T - \lambda_1 I)^{e_1}(w) = 0$.
Therefore, w is a generalized eigenvector of T corresponding to λ_1 but *is not* an eigenvector of T corresponding to λ_1. From a previous exercise, we know that this implies that there exists no basis of V consisting of eigenvectors of T. We've therefore proved the contrapositive of the statement we wanted to prove, thus also proving said statement (the " $\Longleftarrow$ " direction).

Exercise 16
Suppose V is an inner product space and $T \in L(V)$. Suppose

$$a_0 + a_1 z + \ldots + a_{m-1} z^{m-1} + z^m$$

is the minimal polynomial of T. Prove that

$$\overline{a_0} + \overline{a_1} z + \ldots + \overline{a_{m-1}} z^{m-1} + z^m$$

is the minimal polynomial of T^*.

Solution.
Let $p(z)$ be the polynomial given in the exercise. Since p is the minimal polynomial of T, we have that $p(T) = 0$. Therefore, the linear map $p(T) = a_0 I + a_1 T + \ldots + a_{m-1} T^{m-1} + T^m$ is the zero map. By using the properties of the adjoint, we obtain that the adjoint of this linear map is $p_a(T) = \overline{a_0} I + \overline{a_1} T^* + \ldots + \overline{a_{m-1}} (T^*)^{m-1} + (T^*)^m$, where $p_a(z) = \overline{a_0} + \overline{a_1} z + \ldots + \overline{a_{m-1}} z^{m-1} + z^m$. Clearly, $p_a(T^*)$ must also be the zero map. In combination with the fact that p_a is monic, we have two of the conditions for p_a being the minimal polynomial of T^*. We only need to prove that there is no monic polynomial q of degree lower than m such that $q(T^*) = 0$. Suppose that such a q did exist, and that it had the form $q(z) = b_0 + b_1 z + \ldots + b_{n-1} z^{n-1} + z^n, n < m$. By using the fact that $q(T^*) = 0$ and applying again the properties of the adjoint (recall that $(T^*)^* = T$), we conclude that $\overline{b_0} I + \overline{b_1} T + \ldots + \overline{b_{n-1}} T^{n-1} + T^n = 0, n < m$. Because $n < m$, this would mean that p is not the minimal polynomial of T, which is a contradiction. Therefore p_a is indeed the minimal polynomial of T^*.

Exercise 17
Suppose $\mathbf{F} = \mathbf{C}$ and $T \in L(V)$. Suppose the minimal polynomial of T has degree $\dim V$. Prove that the characteristic polynomial of T equals the minimal polynomial of T.

Solution.
We know that the roots of the minimal polynomial p are precisely the eigenvalues of T. Therefore, $p(z) = (z - \lambda_1)^{e_1}(z - \lambda_2)^{e_2} \ldots (z - \lambda_m)^{e_m}$, with λ_i being the eigenvalues of T and e_i some positive integer exponents. The degree of p is $\dim V$, therefore $\sum_i e_i = \dim V$.
The characteristic polynomial of T is $q(z) = (z - \lambda_1)^{d_1}(z - \lambda_2)^{d_2} \ldots (z - \lambda_m)^{d_m}$, with d_i being the multiplicities of the corresponding eigenvalues λ_i. We also know that in complex vector spaces, q must be a polynomial multiple of p, i.e., there exists a polynomial a such that $q(z) = p(z)a(z)$. Observe that q, p have exactly the same zeros. This means that a must be of the form $a(z) = (z - \lambda_1)^{f_1}(z - \lambda_2)^{f_2} \ldots (z - \lambda_m)^{f_m}, f_i \geq 0$. Furthermore, the degree of q also equals $\dim V$. Because $\deg p = \deg q$, we conclude that every f_i must be zero: if this was not the case, the degree of q would be strictly larger than the degree of p, which is a contradiction. Therefore, $a(z) = 1$, which means that $p(z) = q(z)$, i.e. the characteristic polynomial of T equals the minimal polynomial of T.

8.D Jordan Form

Exercise 1
Find the characteristic polynomial and the minimal polynomial of the operator N in Example 8.53.

Solution.
With respect to the basis $z_1 = N^3(z_4), z_2 = N^(2(z_4), z_3 = N(z_4), z_1 = (1,0,0,0)$, the operator N of this example has the following matrix:

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

Observe that 0 is thus the only eigenvalue of N, and N is nilpotent. Thus, the characteristic polynomial must necessarily be $p(z) = z^{\dim V} = z^4$. As for the minimal polynomial, because the characteristic polynomial is a polynomial multiple of it, we need only examine whether T^2 or T^3 are zero. With respect to the above basis, N operates as: $N(z_1) = 0, N(z_2) = z_1, N(z_3) = z_2, N(z_4) = z_3$, thus:

$$N^2(z_1) = 0, N^2(z_2) = N(z_1) = 0, N^2(z_3) = N(z_2) = z_1, N^2(z_4) = N(z_3) = z_2$$

$$N^3(z_1) = 0, N^3(z_2) = N(z_1) = 0, N^3(z_3) = N(N^2(z_2)) = 0, N^3(z_4) = N(z_2) = z_1$$

, therefore neither of these operators are zero, and thus the minimal polynomial must equal the characteristic polynomial.

Exercise 3
Suppose $N \in L(V)$ is nilpotent. Prove that the minimal polynomial of N is z^{m+1}, where m is the length of the longest consecutive string of 1's that appears on the line directly above the diagonal in the matrix of N with respect to any Jordan basis for N.

Solution.
Since N is nilpotent, it holds that $N^{\dim V} = 0$, i.e. $G(0, N) = V$, therefore the characteristic polynomial of N is $p(z) = z^{\dim} V$. We know that the characteristic polynomial is a polynomial multiple of the minimal one. Therefore, the minimal polynomial of N must be of the form $q(z) = z^n$ for some nonnegative integer n. All that remains, therefore, is to show that the minimum value of n for which $N^n = 0$ is $m + 1$, with m described in the exercise.
Consider a Jordan basis for N, that is, a basis of the form

$$N^{m_1}(v_1), \ldots, N(v_1), v_1, N^{m_2}(v_2), \ldots, N(v_2), v_2, \ldots, N^{m_k}(v_k), \ldots N(v_k), v_k$$

, such that $N^{m_1+1}(v_1) = N^{m_2+1}(v_2) = \ldots = N^{m_k+1}(v_k) = 0$. We claim that $m = \max\{m_1, m_2, \ldots, m_k\}$. To see why this is the case, recall that $M(N)$ with respect to this basis is block diagonal, and each of the blocks is a $m_i + 1 \times m_i + 1$ zero matrix, with the exception of the line above the diagonal which is filled with ones. Clearly, inside block i the number of consecutive ones on the line above the diagonal is m_i, therefore the maximum number of such consecutive ones is at least equal to the maximum of $m_1, \ldots, m_k$. If it was larger, this would mean that at least one element that lies on the line above the diagonal *and* does not belong to any block is 1. This is the only way that the ones of at least two blocks, i and $i+1$ are not separated by a zero. However, by the definition of $M(N)$ with respect to the Jordan basis this would imply that $N(N^{m_{i+1}}(v_{i+1})) = v_i \neq 0$, which is of course a contradiction.
If we now raise N to the power $m+1$, we see that for each of the sublists of the basis $N^{m_i}(v_i), \ldots, N(v_i), v_i$, each vector is mapped to $N^x(v_i)$, where $x \geq m + 1 \geq m_i + 1$, thus $N^x(v_i) = 0$. Therefore, N^{m+1} is indeed zero for all vectors of the basis. Now consider any power N^n with $n < m+1$. Then $n \leq \max\{m_1, \ldots, m_k\}$, therefore n is smaller than at least one $m_i + 1$. If we consider the sublist $N^{m_i}(v_i), \ldots, N(v_i), v_i$, then $N^n(v_i)$ with $n < m_i + 1$ appears somewhere in the sublist, and is therefore a vector of the Jordan basis and thus cannot be zero. Thus N^n cannot be zero everywhere.
We've therefore shown that the minimum exponent for which N^n equals zero is $m + 1$, thus proving also that the minimal polynomial of N is $q(z) = z^{m+1}$.

Exercise 4
Suppose $T \in L(V)$ and $v_1, \ldots, v_n$ is a basis of V that is a Jordan basis for T. Describe the matrix of T with respect to the basis $v_n, \ldots, v_1$ obtained by reversing the order of the v's.

Solution.
A Jordan basis for T has the form

$$N^{m_1}(v_1), \ldots, N(v_1), v_1, N^{m_2}(v_2), \ldots, N(v_2), v_2, \ldots, N^{m_k}(v_k), \ldots N(v_k), v_k$$

, such that $N^{m_1+1}(v_1) = \ldots = N^{m_k+1}(v_k) = 0$. Reversing the basis therefore gives us

$$v_k, N(v_k), \ldots, N^{m_k}(v_k), \ldots, v_1, N(v_1), \ldots, N^{m_1}(v_1)$$

.
It is easy to see that with respect to this reversed basis, the matrix of T must equal the matrix that results from taking the block diagonal matrix with respect to the Jordan basis, reversing the order of the blocks on the diagonal and transposing each block. To see this, consider the first $m_k + 1$ columns of the new matrix: they must correspond to the *last* $m_k + 1$ columns of the original one, thus to the last block. Inside the block itself, the original matrix specified that each of the vectors should be mapped to λ_k times the vector plus (for all vectors except the first vector of the sub-list) the immediately preceding vector of the basis. When reversing the order, nothing changes except for the fact that each vector should be mapped to λ_k times the vector plus (for all vectors except the *last* vector of the sub-list) the immediately *following* vector of the sub-list, which leads to a block with λ_k on the diagonal and 1's on the line directly below it, i.e. the transpose of the original block.

Chapter 9

Operators on Real Vector Spaces

9.A Complexification

Exercise 1
Prove 9.3, i.e., that if V is a real vector space, under the definitions given in 9.2, $V_{\mathbf{C}}$ is a complex vector space.

Solution. Let us examine one by one the properties that vector addition and scalar multiplication must have in order for a set to be a vector space.

- **Commutativity of addition**: let $u, v \in V_{\mathbf{C}}$. This means that $u = u_1 + u_2 i, v = v_1 + v_2 i, u_1, u_2, v_1, v_2 \in V$. Then, by the definition of addition in $V_{\mathbf{C}}$, we have that: $u + v = (u_1 + v_1) + (v_1 + v_2)i, v + u = (v_1 + u_1) + (v_2 + u_2)i$. Since V is a vector space, by the commutativity of its respective addition operation we have that $u_1 + v_1 = v_1 + u_1, u_2 + v_2 = v_2 + u_2$. Therefore, viewed as ordered pairs, $u + v, v + u$'s respective coordinates are equal, thus they themselves are equal, i.e. addition on $V_{\mathbf{C}}$ is indeed commutative.

- **Associativity of addition**: let $u, v, w \in V_{\mathbf{C}}$, i.e. $u = u_1 + u_2 i, v = v_1 + v_2 i, w = w_1 + w_2 i$. Then:

$$(u + v) + w = ((u_1 + v_1) + (u_2 + v_2)i) + w_1 + w_2 i = (u_1 + v_1 + w_1) + (u_2 + v_2 + w_2)i$$

$$u + (v + w) = u_1 + u_2 i + ((v_1 + w_1) + (v_2 + w_2)i) = (u_1 + v_1 + w_1) + (u_2 + v_2 + w_2)i$$

  , where $u_1 + v_1 + w_1, u_2 + v_2 + w_2$ are well-defined due to the associativity of addition on V. The above two equalities imply that $(u + v) + w = u + (v + w)$, therefore addition on $V_{\mathbf{C}}$ is indeed associative.

- **Additive identity**: Consider the element $0_{\mathbf{C}} = 0 + 0i$, 0 being the additive identity of V. Then, for any $v \in V_{\mathbf{C}}$, we have that $v + 0_{\mathbf{C}} = v_1 + v_2 i + 0 + 0i = (v_1 + 0) + (v_2 + 0)i = v_1 + v_2 i = v$, where we used the fact that 0 is the additive identity on V. Thus, $0_{\mathbf{C}}$ is indeed the/an additive identity on $V_{\mathbf{C}}$.

- **Additive inverse**: Let $v = v_1 + v_2 i$ be an element of $V_{\mathbf{C}}$. Then, let $w = (-v_1) + (-v_2)i$, where $-v_1, -v_2$ are the additive inverses of $v_1, v_2 \in V$. We have, then, that:

$$v + w = (v_1 + (-v_1)) + (v_2 + (-v_2))i = 0 + 0i = 0_{\mathbf{C}}$$

  , by the definition of additive inverses in V. Since we've shown that $0_{\mathbf{C}}$ is the additive identity of $V_{\mathbf{C}}$, we have proved that for any given $v \in V_{\mathbf{C}}$ there exists (in $V_{\mathbf{C}}$) an additive inverse of v.

- **Multiplicative identity**: For any $v \in V_{\mathbf{C}}$, and for $1 \in \mathbf{C}$ we have that $1v = 1(v_1 + v_2 i) = (1v_1) + (1v_2)i = v_1 + v_2 i$, since it is also true that $1 \in \mathbf{R}$, and 1 is the multiplicative identity on V.

- **Associativity of scalar multiplication**: For $a, b \in \mathbf{C}$, and $v \in V_{\mathbf{C}}$ we have that $a = a_1 + a_2 i, b = b_1 + b_2 i, v = v_1 + v_2 i$. Then:

$$(ab)v = ((a_1 + a_2 i)(b_1 + b_2 i))v = (a_1 b_1 - a_2 b_2 + (a_1 b_2 + a_2 b_1)i)v =$$

$$((a_1b_1 - a_2b_2)v_1 - (a_1b_2 + a_2b_1)v_2) + ((a_1b_1 - a_2b_2)v_2 + (a_1b_2 + a_2b_1)v_1)i$$

And:

$$a(bv) = a((b_1 + b_2i)(v_1 + v_2i)) = a((b_1v_1 - b_2v_2) + (b_1v_2 + b_2v_1)i) =$$

$$(a_1 + a_2i)((b_1v_1 - b_2v_2) + (b_1v_2 + b_2v_1)i) =$$

$$(a_1b_1v_1 - a_1b_2v_2 - a_2b_1v_2 - a_2b_2v_1) + (a_2b_1v_1 - a_2b_2v_2 + a_1b_1v_2 + a_1b_2v_1)i =$$

$$((a_1b_1 - a_2b_2)v_1 - (a_1b_2 + a_2b_1)v_2) + ((a_2b_1 + a_1b_2)v_1 + (a_1b_2 - a_2b_2)v_2)i$$

, where we use the properties of complex number multiplication, the distributive properties on V and the definition of scalar multiplication on $V_{\mathbf{C}}$. We observe that the two resulting quantities above are equal, thus indeed $(ab)v = a(bv)$.

- Distributive properties: the proofs here are done as in the case for associativity of scalar multiplication, and are omitted for the sake of brevity.

Therefore, $V_{\mathbf{C}}$ is indeed a -complex- vector space.

Exercise 2
Verify that if V is a real vector space and $T \in L(V)$, then $T_{\mathbf{C}} \in L(V_{\mathbf{C}})$.

Solution.
We need to prove three things. One, that for any $v \in V_{\mathbf{C}}$, $T_{\mathbf{C}}(v) \in V_{\mathbf{C}}$. Two, that $T_{\mathbf{C}}$ has the additivity property and three that it has the homogeneity property.
Then, let $v \in V_{\mathbf{C}}$, therefore $v = v_1 + v_2i$ for $v_1, v_2 \in V$. By definition, $T_{\mathbf{C}}(v) = T(v_1) + T(v_2)i$. Because T is an operator, $T(v_1), T(v_2) \in V$. Thus, by definition of the complexification of V, $T(v_1) + T(v_2)i \in V_{\mathbf{C}}$. Therefore $T_{\mathbf{C}}$ is an operator on $V_{\mathbf{C}}$.
Let then also $w \in V_{\mathbf{C}}, w = w_1 + w_2i, w_1, w_2 \in V$. Then

$$T_{\mathbf{C}}(v + w) = T_{\mathbf{C}}((v_1 + w_1) + (v_2 + w_2)i) =$$

$$T(v_1 + w_1) + T(v_2 + w_2)i = T(v_1) + T(w_1) + (T(v_2) + T(w_2))i =$$

$$(T(v_1) + T(v_2)i) + (T(w_1) + T(w_2)i) = T_{\mathbf{C}}(v) + T_{\mathbf{C}}(w)$$

, by the additivity of T and the definition of addition on $V_{\mathbf{C}}$. Therefore, $T_{\mathbf{C}}$ has the additivity property.
Let also $\lambda = a + b_i \in \mathbf{C}$. Then:

$$T_{\mathbf{C}}(\lambda v) = T_{\mathbf{C}}((a + bi)(v_1 + v_2)i) = T_{\mathbf{C}}(av_1 - bv_2 + (bv_1 + av_2)i)$$

$$= T(av_1 - bv_2) + T(bv_1 + av_2)i = aT(v_1) - bT(v_2) + (bT(v_1) + aT(v_2))i$$

Also:

$$\lambda T_{\mathbf{C}}(v) = (a + bi)(T(v_1) + T(v_2)i) = aT(v_1) - bT(v_2) + (T(v_1)b + T(v_2)a)i$$

, where in both cases we used the homogeneity of T and the definition of scalar multiplication on $V_{\mathbf{C}}$. Since the two quantities are equal, we have $T_{\mathbf{C}}(\lambda v) = \lambda T_{\mathbf{C}}(v)$, therefore $T_{\mathbf{C}}$ has the homogeneity property, thus completing the proof that $T_{\mathbf{C}} \in L(V_{\mathbf{C}})$.

Exercise 3
Suppose V is a real vector space and $v_1, \ldots, v_m \in V$. Prove that $v_1, \ldots, v_m$ is linearly independent in $V_{\mathbf{C}}$ if and only if $v_1, \ldots, v_m$ is linearly independent in V.

Solution.
$\Longrightarrow$: Suppose $v_1, \ldots, v_m$ is linearly independent in $V_{\mathbf{C}}$. Since $v_i \in V$, we think of them as elements of $V_{\mathbf{C}}$ by identifying each of them with $v_i + 0i \in V_{\mathbf{C}}$ (this mapping is an isomorphism). Suppose now that for some $a_j \in \mathbf{R}$, it holds that $\sum_j a_j v_j = 0$. Then, we have that:

$$\sum_j a_j v_j = 0 \implies \sum_j a_j(v_j + 0i) = 0_{\mathbf{C}}$$

, since the zero vector of $V_{\mathbf{C}}$ has zero as both its real and imaginary "part". Since $v_1, \ldots, v_m$ are linearly independent in $V_{\mathbf{C}}$, the last equality implies that $a_i = 0$ for all i, thus $v_1, \ldots, v_m$ are also linearly independent in V.
$\Longleftarrow$: Suppose now that $v_1, \ldots, v_m$ are linearly independent in V. We want to examine whether they are linearly independent in $V_{\mathbf{C}}$, i.e. whether $v_1 + 0i, v_2 + 0i, \ldots, v_m + 0i$ are linearly independent in $V_{\mathbf{C}}$. Suppose that for some $a_j \in \mathbf{C}$, $\sum_j a_j(v_j + 0i) = 0_{\mathbf{C}}$. This implies that:

$$\sum_j a_j(v_j + 0i) = 0_{\mathbf{C}} \implies \sum_j (Re\{a_j\} + Im\{a_j\}i)(v_j + 0_j) = 0_{\mathbf{C}} \implies \sum_j Re\{a_j\}v_j + \sum_j Im\{j\}v_j i = 0_{\mathbf{C}}$$

Due to the definition of the zero vector of $V_{\mathbf{C}}$, both the real and imaginary "parts" of the left-hand side must be zero. This fact, combined with the linear independence of $v_1, \ldots, v_m$ in V and the fact that all $Re\{a_j\}, Im\{a_j\}$ are real numbers implies that all of them are zero, hence that each a_j is also zero and therefore that $v_1, \ldots, v_m$ are linearly independent in $V_{\mathbf{C}}$ as well.

Exercise 4
Suppose V is a real vector space and $v_1, \ldots, v_m \in V$. Prove that $v_1, \ldots, v_m$ spans $V_{\mathbf{C}}$ if and only if $v_1, \ldots, v_m$ spans V.

Solution.
$\Longrightarrow$: Suppose that $v_1, \ldots, v_m$ spans $V_{\mathbf{C}}$. Suppose also that $v \in V$. We can identify v with the element $v + 0i$ of $V_{\mathbf{C}}$. Since v_j spans $V_{\mathbf{C}}$, there exist $a_j \in \mathbf{C}$ such that:

$$\sum_j a_j(v_j + 0i) = v + 0i \implies \sum_j (Re\{a_j\} + Im\{a_j\}i)(v_j + 0i) = v + 0i$$

$$\implies \sum_j Re\{a_j\}v_j + \sum Im\{a_j\}v_j i = v + 0i$$

Equality of elements of $V_{\mathbf{C}}$ implies equality of both their real and imaginary "parts". Thus we have that:

$$\sum_j Re\{a_j\}v_j = v$$

, therefore v can be written as a linear combination of v_j, and since it was arbitrarily selected in V we conclude that $v_1, \ldots, v_m$ spans V.
$\Longleftarrow$: Suppose that $v_1, \ldots, v_m$ spans V. Suppose also that $w \in V_{\mathbf{C}}$. Then there exist $u, v \in V$ such that $w = u + vi$. Additionally, since $v_1, \ldots, v_m$ spans V there exist $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbf{R}$ such that:

$$\sum_j a_j v_j = u, \sum_j b_j v_j = v$$

This implies that:

$$w = u + vi = \sum_j a_j v_j + \sum_j b_j v_j i = \sum_j (a_j + b_j i)(v_j + 0i)$$

Since $a_j + b_j i \in \mathbf{C}$, we have written w as a linear combination of v_j, and since w was arbitrarily selected in $V_{\mathbf{C}}$, $v_1, \ldots, v_m$ spans $V_{\mathbf{C}}$.

Exercise 5
Suppose that V is a real vector space and $S, T \in L(V)$. Show that $(S + T)_{\mathbf{C}} = S_{\mathbf{C}} + T_{\mathbf{C}}$ and that $(\lambda T)_{\mathbf{C}} = \lambda T_{\mathbf{C}}$ for every $\lambda \in \mathbf{R}$.

Solution.
Let $w \in V_{\mathbf{C}}$. We have then that there exist $u, v \in V$ such that $w = u + vi$, and that:

$$(S+T)_{\mathbf{C}}(w) = (S+T)_{\mathbf{C}}(u+vi) = (S+T)(u) + (S+T)(v)i = S(u) + T(u) + (S(v) + T(v))i =$$
$$(S(u) + S(v)i) + (T(u) + T(v)i) = S_{\mathbf{C}}(w) + T_{\mathbf{C}}(w)$$

, by the definition of the complexification of an operator, the definition of addition in $L(V)$ and by the definition of addition in the complexification of V. Thus $(S + T)_{\mathbf{C}} = S_{\mathbf{C}} + T_{\mathbf{C}}$.
Similarly, for $\lambda \in \mathbf{R}$, the operator λT: is well-defined and belongs in $L(V)$, thus its complexification is also well-defined and we have that:

$$(\lambda T)_{\mathbf{C}}(w) = (\lambda T)_{\mathbf{C}}(u+vi) = (\lambda T)(u) + (\lambda T)(v)i = \lambda T(u) + \lambda T(v)i = \lambda(T(u) + T(v)i) = \lambda T_{\mathbf{C}}(w)$$

, where we used the definition of the complexification of an operator and the definition of scalar multiplication in $L(V)$ and $V_{\mathbf{C}}$. Therefore, it is indeed true that $(\lambda T)_{\mathbf{C}} = \lambda T_{\mathbf{C}}$ for $\lambda \in \mathbf{R}$.

Exercise 6
Suppose V is a real vector space and $T \in L(V)$. Prove that $T_{\mathbf{C}}$ is invertible if and only if T is invertible.

Solution.
$\implies$: Suppose $T_{\mathbf{C}}$ is invertible. Because $T, T_{\mathbf{C}}$ are operators, to show that T is invertible it suffices to show that it is injective or surjective. Suppose that for some $v \in V$, $T(V) = 0$. Then, we have that:

$$T_{\mathbf{C}}(v + 0i) = T(v) + T(0)i = 0 + 0i = 0_{\mathbf{C}}$$

Because $T_{\mathbf{C}}$ is invertible, it is injective. Thus, the above equation can only hold if $v + 0i = 0_{\mathbf{C}}$, which implies that $v = 0$, which implies that T is indeed injective, and thus invertible.
$\impliedby$: Suppose now that T is invertible. Suppose that for some $w \in V_{\mathbf{C}}, w = u + vi, u, v \in V$ it holds that $T_{\mathbf{C}}(w) = 0_{\mathbf{C}}$. Then, we have that:

$$T_{\mathbf{C}}(w) = 0_{\mathbf{C}} \implies T(u) + T(v)i = 0 + 0i$$

This equality implies that $T(u) = 0, T(v) = 0$. Since T is injective, this can only happen for $u = v = 0$, which yields $w = 0 + 0i = 0_{\mathbf{C}}$, which implies $T_{\mathbf{C}}$ is injective, and thus invertible.

Exercise 7
Suppose V is a real vector space and $N \in L(V)$. Prove that $N_{\mathbf{C}}$ is nilpotent if and only if N is nilpotent.

Solution.
Previously in the book we have seen that $T^k_{\mathbf{C}}(u+vi) = T^k(u)+T^k(v)i$. Based on this observation, we have that:
$\implies$: Suppose $N_{\mathbf{C}}$ is nilpotent. Let $v \in V$. Then we have that:

$$N_{\mathbf{C}}^{\dim V_{\mathbf{C}}}(v+0i)=0_{\mathbf{C}} \implies N^{\dim V_{\mathbf{C}}}(v)+N^{\dim V_{\mathbf{C}}}(0)i=0_{\mathbf{C}} \implies N^{\dim V}(v)+0i=0_{\mathbf{C}}$$

, where we used the equality of the dimensions of $V, V_{\mathbf{C}}$. This implies that for any v, $N^{\dim V}(v)=0$, therefore N is indeed nilpotent.

$\impliedby$: Suppose N is nilpotent. Let $w=u+vi, w\in V_{\mathbf{C}}, u,v\in V$. Because N is nilpotent, it holds that:

$$N^{\dim V}(u)=0, N^{\dim V}(v)=0 \implies N^{\dim V_{\mathbf{C}}}(u)=0, N^{\dim V_{\mathbf{C}}}(v)=0$$

Then we have that:

$$N_{\mathbf{C}}^{\dim V_{\mathbf{C}}}(w)=N^{\dim V_{\mathbf{C}}}(u)+N^{\dim V_{\mathbf{C}}}(v)i=0+0i=0_{\mathbf{C}}$$

Therefore, for any $w\in V_{\mathbf{C}}, N_{\mathbf{C}}^{\dim V_{\mathbf{C}}}(w)=0$, which means that $N_{\mathbf{C}}$ is nilpotent.

Exercise 16
Suppose V is a real vector space. Prove that there exists $T\in L(V)$ such that $T^2=-I$ if and only if V has even dimension.

Solution.
$\implies$: Suppose that there exists a $T\in L(V)$ such that $T^2=-I$. Suppose then that V has odd dimension. We know then that T has at least one eigenvalue $\lambda\in\mathbf{R}$, and by definition at least one non-zero eigenvector v corresponds to this eigenvalue. It would then hold that:

$$T^2(v)=-I(v) \implies T(T(v))=-v \implies T(\lambda v)=-v \implies \lambda^2 v+v=0 \implies (\lambda^2+1)v=0$$

Since V is non-zero, this implies that $\lambda^2=-1$ which is a contradiction because $\lambda\in\mathbf{R}$. Therefore, the dimension of V has to be even.

$\impliedby$: Suppose now that the dimension of a real vector space V is even. This mean that any basis of it is of the form $v_1, v_2, \dots, v_{2k}$. Pick one such basis and consider the following T, defined by its values on this basis:

$$T(v_i)=-v_{i+1}, \text{for } i \text{ odd}, T(v_i)=v_{i-1}, \text{ for } i \text{ even}$$

Observe that because the dimension is even, for any v_i with i odd and at most equal to $2k-1$, v_{i+1} is well-defined as a vector of the basis. Symmetrically, for any v_i with i even and at at most equal to $2k$, v_{i-1} is well-defined as a vector of the basis. As a consequence of this, we observe that:

$$T^2(v_i)=T(T(v_i))=T(-v_{i+1})=-T(v_{i+1})=-v_i, \text{for } i \text{ odd}$$

$$T^2(v_i)=T(T(v_i))=T(v_{i-1})=-v_i, \text{for } i \text{ even}$$

It is therefore true that for this $T, T^2=-I$, thus completing the proof in the other direction.

Exercise 18
Suppose V is a real vector space and $T\in L(V)$. Prove that the following are equivalent.
(a) All the eigenvalues of $T_{\mathbf{C}}$ are real.
(b) There exists a basis of V with respect to which T has an upper-triangular matrix.
(c) There exists a basis of V consisting of generalized eigenvectors of T.

Solution.
(c) $\implies$ (b): If there exists a basis of V consisting of generalized eigenvectors of T, we know from previous results that the matrix of T with respect to this basis is upper triangular.
(b) $\implies$ (a): Suppose that there exists a basis of V with respect to which T has an upper-triangular matrix. Again from previous results, we know that the eigenvalues of T are precisely the elements of the diagonal of this matrix, and that the multiplicity of each one equals the number of times it appears on this diagonal. We also know that each eigenvalue of T is an eigenvalue of $T_{\mathbf{C}}$, and also that $\dim V=\dim V_{\mathbf{C}}$.

From these facts we conclude first that the sum of the multiplicities of the eigenvalues of T equals $\dim V$, since the number of diagonal elements of $M(T)$ is of course $\dim V$. Therefore, it also equals $\dim V_{\mathbf{C}}$. This means, however, that these must be *all* of the eigenvalues of $T_{\mathbf{C}}$, since in a complex vector space the sum of the multiplicities of the eigenvalues of an operator always equals the dimension of the space. Therefore, $T_{\mathbf{C}}$ cannot have any eigenvalues that are non-real, therefore all of its eigenvalues are real.
(a) $\implies$ (c): Suppose that all of the eigenvalues of $T_{\mathbf{C}}$ are real. Pick any generalized eigenvector $w \in V_mathbfC, w = u + vi, u, v \in V$ corresponding to an eigenvalue $\lambda \in \mathbf{R}$. We then have that:

$$(T_{\mathbf{C}} - \lambda I_{\mathbf{C}})^{\dim V_{\mathbf{C}}}(w) = 0_{\mathbf{C}} \implies (T - \lambda I)^{\dim V}(u) + (T - \lambda I)^{\dim V}(v)i = 0 + 0i$$
$$\implies (T - \lambda I)^{\dim V}(u) = 0, (T - \lambda I)^{\dim V}(v) = 0$$

, where we used a previously observed fact about powers of the complexification of an operator, the fact that $\lambda \in \mathbf{R}$ and the definition of the zero vector of $V_{\mathbf{C}}$. We therefore conclude that u, v are generalized eigenvectors of T corresponding to λ. We can always find a basis of $V_{\mathbf{C}}$ consisting of generalized eigenvectors $w_j = u_j + v_j i$ of $T_{\mathbf{C}}$. Therefore, for any element $v \in V$, we can write $v + 0i$ as a linear combination of those vectors. This also implies that v can be written as a linear combination (with real coefficients) of all of the u_j, v_j. Since we showed that all of those are generalized eigenvectors of T, they form a spanning list of V consisting of generalized eigenvectors of T.
Clearly, we can extract a basis of V from this list, which will also consist of generalized eigenvectors of T, thus completing the proof.

9.B Operators on Real Inner Product Spaces

Exercise 2
Prove that every isometry on an odd-dimensional real inner product space has 1 or -1 as an eigenvalue.

Solution.
Every operator S on a real vector space of odd dimension has an eigenvalue. Suppose now that S is an isometry, λ is said eigenvalue and $v \neq 0$ a corresponding eigenvector. Then we have that:

$$||Sv|| = ||v|| \implies ||\lambda v|| = ||v|| = \implies |\lambda| = 1$$

, since v has a non-zero norm. Therefore, since $\lambda \in \mathbf{R}$, it must hold that $\lambda = 1$ or $\lambda = -1$.

Exercise 3
Suppose V is a real inner product space. Show that

$$\langle u + iv, x + iy \rangle = \langle u, x \rangle + \langle v, y \rangle + (\langle v, x \rangle - \langle u, v \rangle)i$$

for $u, v, x, y \in V$ defines a complex inner product on $V_{\mathbf{C}}$.

Solution.
Let us examine one by one the necessary and sufficient properties for a complex inner product:

- **Positivity**: We need to examine whether $\langle u + iv, u + iv \rangle \geq 0$ for every $u, v \in V$. We have that:

  $$\langle u+iv, u+iv \rangle = \langle u, u \rangle + \langle v, v \rangle + (\langle v, u \rangle - \langle u, v \rangle)i = \langle u, u \rangle + \langle v, v \rangle + (\langle v, u \rangle - \langle v, u \rangle)i = \langle u, u \rangle + \langle v, v \rangle \geq 0$$

  , due to the symmetry and positivity of the real inner product on V.

- **Definiteness**: From the equality derived above, we have that $\langle u + iv, u + iv \rangle - \langle u, u \rangle + \langle v, v \rangle$. This is a sum of nonnegative terms that thus only becomes zero when both terms are zero. Due to the definiteness of the real inner product on V, this happens precisely when $u = v = 0$, which yields $u + iv = 0_{\mathbf{C}}$. Therefore, $\langle u + iv, u + iv \rangle$ is zero iff $u + iv = 0_{\mathbf{C}}$, i.e. the defined (complex) function has the definiteness property.

- **Additivity in the first slot**: Let now $w_1 = u_1 + v_1 i, w_2 = u_2 + v_2 i, w_3 = u_3 + v_3 i, w_1, w_2, w_3 \in V_{\mathbf{C}}, u_1, u_2, u_3, v_1, v_2, v_3 \in V$. Then:

$$\langle w_1+w_2, w_3\rangle = \langle (u_1+u_2)+(v_1+v_2)i, u_3+v_3 i\rangle = \langle u_1+u_2, u_3\rangle+\langle v_1+v_2, v_3\rangle+(\langle v_1+v_2, u_3\rangle-\langle u_1+u_2, v_3\rangle)i$$

$$= \langle u_1, u_3\rangle + \langle u_2, u_3\rangle + \langle v_1, v_3\rangle + \langle v_2, v_3\rangle + (\langle v_1, u_3\rangle + \langle v_2, u_3\rangle - \langle u_1, v_3\rangle - \langle u_2, v_3\rangle)i$$

$$= (\langle u_1, u_3\rangle + \langle v_1, v_3\rangle + (\langle v_1, u_3\rangle - \langle u_1, u_3\rangle)i) + (\langle u_2, u_3\rangle + \langle v_2, v_3\rangle + (\langle v_2, u_3\rangle - \langle u_2, v_3\rangle)i)$$

$$= \langle w_1, w_3\rangle + \langle w_2, w_3\rangle$$

, where we used the first-slot linearity of the real inner product on V as well as the properties of complex numbers. Therefore, the complex function is linear in the first slot as well.

- **Homogeneity in the first slot**: Let $\lambda = a+bi \in \mathbf{C}$ and $w_1 = u_1+v_1 i, w_2 = u_2+v_2 i \in V_{\mathbf{C}}, u_1, u_2, v_1, v_2 \in V$. Then:

$$\langle \lambda w_1, w_2\rangle = \langle (a+bi)(u_1+v_1 i), u_2+v_2 i\rangle = \langle (au_1 - bv_1) + (av_1 + bu_1)i, u_2 + v_2 i\rangle$$

$$= \langle au_1 - bv_1, u_2\rangle + \langle av_1 + bu_1, v_2\rangle + (\langle av_1 + bu_1, u_2\rangle - \langle au_1 - bv_1, v_2\rangle)i$$

Additionally:

$$\lambda\langle w_1, w_2\rangle = (a+bi)(\langle u_1, u_2\rangle + \langle v_1, v_2\rangle + (\langle v_1, u_2\rangle - \langle u_1, v_2\rangle)i$$

$$= a(\langle u_1, u_2\rangle + \langle v_1, v_2\rangle) - b(\langle v_1, u_2\rangle - \langle u_1, v_2\rangle) + (a(\langle v_1, u_2\rangle - \langle u_1, v_2\rangle) + b(\langle u_1, u_2\rangle + \langle v_1, v_2\rangle))i$$

$$= \langle au_1 - bv_1, u_2\rangle + \langle av_1 + bu_1, v_2\rangle + (\langle av_1 + bu_1, u_2\rangle - \langle au_1 - bv_1, v_2\rangle)i$$

, where we used the first-slot homogeneity of the real inner product and the definition of complex multiplication. Notice that we've arrived at the same expression both times, thus $\langle \lambda w_1, w_2\rangle = \lambda\langle w_1, w_2\rangle$, which means we've proved first-slot homogeneity for the complex function.

- **Conjugate symmetry**: Let $w_1 = u_1 + v_1 i, w_2 = u_2 + v_2 i, w_1, w_2 \in V_{\mathbf{C}}, u_1, u_2, v_1, v_2 \in V$. We have that:

$$\langle w_2, w_1\rangle = \langle u_2 + v_2 i, u_1 + v_1 i\rangle = \langle u_2, u_1\rangle + \langle v_2, v_1\rangle + (\langle v_2, u_1\rangle - \langle u_2, v_1\rangle)i$$

$$= \langle u_1, u_2\rangle + \langle v_1, v_2\rangle - (\langle v_1, u_2\rangle - \langle u_1, v_2\rangle)i = \overline{\langle u_1, u_2\rangle + \langle v_1, v_2\rangle + (\langle v_1, u_2\rangle - \langle u_1, v_2\rangle)i}$$

$$= \overline{\langle w_1, w_2\rangle}$$

, where we used the symmetry of the real inner product on V and the definition of complex conjugates.

Therefore, we've proved all of the properties necessary for a function to be a complex inner product on $V_{\mathbf{C}}$.

> **Exercise 4**
> Suppose V is a real inner product space and $T \in L(V)$ is self-adjoint. Show that $T_{\mathbf{C}}$ is a self-adjoint operator on the inner product space $V_{\mathbf{C}}$ defined by the previous exercise.

Solution.
To prove that $T_{\mathbf{C}}$ is self-adjoint, we need to prove that for any two $w_1, w_2 \in V_{\mathbf{C}}$, $\langle T_{\mathbf{C}} w_1, w_2\rangle = \langle w_1, T_{\mathbf{C}} w_2\rangle$, where the inner product denotes the one defined in the previous exercise. We have that:

$$\langle T_{\mathbf{C}} w_1, w_2\rangle = \langle T_{\mathbf{C}}(u_1 + v_1 i), u_2 + v_2 i\rangle = \langle T(u_1) + T(v_1)i, u_2 + v_2 i\rangle$$

$$= \langle T(u_1), v_2\rangle + \langle T(v_1), v_2\rangle + (\langle T(v_1), u_2\rangle - \langle T(u_1), v_2\rangle)i$$

And:

$$\langle w_1, T_{\mathbf{C}} w_2 \rangle = \langle u_1 + v_1 i, T(u_2) + T(v_2) i \rangle$$
$$= \langle u_1, T(u_2) \rangle + \langle v_1, T(u_2) \rangle + (\langle v_1, T(u_2) \rangle - \langle u_1, T(v_2) \rangle) i$$
$$= \langle T(u_1), u_2 \rangle + \langle T(v_1), u_2 \rangle + (\langle T(v_1), u_2 \rangle - \langle T(u_1), v_2 \rangle) i$$

, where we used the fact that T is self-adjoint in V. Observe that the two expressions end up being equal, so it is indeed the case that $\langle T_{\mathbf{C}} w_1, w_2 \rangle = \langle w_1, T_{\mathbf{C}} w_2 \rangle$, thus $T_{\mathbf{C}}$ is self-adjoint.

Exercise 5
Use the previous exercise to give a proof of the Real Spectral Theorem (7.29) via complexification and the Complex Spectral Theorem (7.24).

Solution.
Suppose first that an operator T in a real vector space V has a diagonal matrix with respect to some orthonormal basis of V. We know then that the matrix of T^* with respect to this basis equals the conjugate transpose of the matrix of T, which due to being diagonal and real equals itself. Therefore T is self-adjoint.
In the other direction, suppose that T is self-adjoint. By exercise 4, $T_{\mathbf{C}}$ is also self-adjoint. Therefore it is normal, and by the Complex Spectral Theorem, there exists a basis $e_1 + f_1 i, e_2 + f_2 i, \ldots, e_n + f_n i$ of $V_{\mathbf{C}}$ with respect to which $M(T)$ is diagonal. More specifically, because $T_{\mathbf{C}}$ is self-adjoint, $M(T)$ must equal its conjugate transpose, and thus all entries on the diagonal are real. Let us call them $a_1, a_2, \ldots, a_n$, and note that they equal precisely the eigenvalues of $T_{\mathbf{C}}$. It is therefore true that:

$$T_{\mathbf{C}}(e_j + f_j i) = T(e_j) + T(f_j) i = a_j (e_j + f_j i)$$

Therefore, $T(e_j) = a_j e_j, T(f_j) = a_j f_j$, i.e. e_j, f_j are eigenvectors of T. By taking all of e_j, f_j we form a spanning list for V since $e_j + i f_j$ are a basis of $V_{\mathbf{C}}$. From this we can extract a basis of V consisting of eigenvectors of T.
Because T is self-adjoint, it is normal, and because it is normal, eigenvectors corresponding to distinct eigenvalues are orthogonal. Because V has a basis consisting of eigenvectors of T, it can be written as a direct sum of all $E(\lambda_i, T)$, λ_i being the discrete eigenvalues of T. Form an orthonormal basis of each $E(\lambda_i, T)$ by applying the Gram-Schmidt procedure to the eigenvectors corresponding to λ_i. By the previous observation that eigenvectors corresponding to distinct eigenvalues of T are orthogonal, if we concatenate all of these bases we obtain an orthonormal basis of V consisting of eigenvectors of T. With respect to this basis it is of course true that T has a diagonal matrix, therefore proving the other direction of the Real Spectral Theorem.

Chapter 10

Trace and Determinant

10.A Trace

Exercise 1
Suppose $T \in L(V)$ and $v_1, \ldots, v_n$ is a basis of V. Prove that the matrix $M(T, (v_1, \ldots, v_n))$ is invertible if and only if T is invertible.

Solution.
$\Longrightarrow$: Suppose that $A = M(T, (v_1, \ldots, v_n))$ is an invertible matrix. Then there exists some matrix A^{-1} such that $AA^{-1} = I$. We know that $L(V)$ is isomorphic to the vector space of square matrices of size $\dim V \times \dim V$ with respect to this basis. That means that there exists an operator $S \in L(V)$ such that $M(S, (v_1, \ldots, v_n) = A^{-1}$. Let $R_1 = TS$. Observe that

$$M(R, (v_1, \ldots, v_n)) = M(T, (v_1, \ldots, v_n))M(S, (v_1, \ldots, v_n)) = AA^{-1} = M(I, (v_1, \ldots, v_n))$$

. Again, due to the two vector spaces being isomorphic, equality of matrices implies equality of operators, i.e. $R_1 = TS = I$. In the same way we can prove that $ST = I$, hence S is the inverse of T, hence the inverse of T does exist.
$\Longleftarrow$: If T is invertible and $S = T^{-1}$, then let $A = M(T, (v_1, \ldots, v_n)), B = M(S, (v_1, \ldots, v_n))$. Then $M(I, (v_1, \ldots, v_n)) = M(TS, (v_1, \ldots, v_n)) = M(T, (v_1, \ldots, v_n))M(S, (v_1, \ldots, v_n)) \implies I = AB$. Similarly $BA = I$, which by definition of the inverse of a matrix means that $B = A^{-1}$, i.e. A is invertible.

Exercise 2
Suppose A and B are square matrices of the same size and $AB = I$. Prove that $BA = I$.

Solution.
Let n be the "dimension" of A, B, i.e. A, B are square matrices of size $n \times n$. Then if $V = \mathbf{F}^n$, if we think of A, B as matrices with respect to the standard basis of V, we know that $L(V)$ and the vector space of those matrices are isomorphic. This means that there exist unique operators T, S such that $M(T) = A, M(S) = B$ (with respect to the standard basis, omitted for brevity). It is true that $AB = I$. This means that $M(TS) = M(T)M(S) = AB = I = M(I)$. Again due to the isomorphism between $L(V)$ and the vector space of matrices, we conclude that $TS = I$. We claim that S is invertible and that $S^-1 = T$, which would directly imply that $M(ST) = M(I) = I = M(S)M(T) = BA$. To prove this we will prove that S is injective. Suppose it was not, in which case there exists a non-zero v such that $S(v) = 0$. Then $TS(v) = T(0) = 0 = I(v) = v$ which is a contradiction. Therefore S is injective and as a consequence invertible.
Therefore, if $v_1, \ldots, v_n$ is a basis of V, $S(v_1), \ldots S(v_n)$ is also a basis of V. It must hold that $T(S(v_i)) = I(v_i) = v_i$. Consider then the values of ST on the basis $S(v_i)$: $ST(S(v_i)) = S(v_i) = I(S(v_i))$, therefore $ST = I$, and because $TS = I$ by definition $T = S^-1$, thus proving the exercise by the observation made above.

Exercise 3
Suppose $T \in L(V)$ has the same matrix with respect to every basis of V. Prove that T is a scalar multiple of the identity operator.

Solution.
To begin, note that if $\dim V = 1$, it is obvious that the only operators that can be defined are scalar multiples of the identity operator, since the basis has exactly one vector which can only be mapped to a scalar multiple of itself. Also, if $\dim V = 0$ it is trivially true that the only operator that can be defined is $T = 0$. Therefore from now on we can assume that $\dim V \geq 2$. Let therefore $v_1, v_2, \ldots, v_n$ be a basis of V, $n \geq 2$. With respect to any basis of V, the matrix of T is the same, so let us call it A. In particular, for any two $i, j, i \neq j$ it holds that:

$$T(v_i) = A_{ii}v_i + A_{ji}v_j + \sum_{k \neq i,j} A_{ki}v_k$$

Consider now the basis that results from exchanging v_i, v_j in the original basis and keeping everything else the same. The matrix of T with respect to it will again equal A, thus for the same i, j we see that we now have:

$$T(v_i) = A_{jj}v_i + A_{ij}v_j + \sum_{k \neq i,j} A_{kj}v_k$$

The reasoning is that the coefficients for $T(v_i)$ now come from the j-th column of the matrix instead of the i-th, and that its j-th row now corresponds to v_i instead of v_j. By equating the two right-hand sides above and observing that these vectors are linearly independent we obtain that $A_{ii} = A_{jj}, A_{ij} = A_{ji}$ and $A_{ki} = A_{kj}$ for all $k \neq i, j$. Now, since we can do this for any two i, j we conclude that all of the diagonal elements of A are equal to some number a and that all off-diagonal elements of A are equal to some number b. The first of these two observations is straightforward, while the second arises from the fact that this is a symmetric matrix whose columns are equal element-wise.
It suffices therefore to show that $b = 0$. Indeed, consider the basis $-v_1, v_2, \ldots, v_n$. Then:

$$T(-v_1) = a(-v_1) + b\sum_{i \neq 1} v_i \implies T(v_1) = av_1 - b\sum_{i \neq 1} v_i$$

However, from our first equation we obtain that:

$$T(v_1) = av_1 + b\sum_{i \neq 1} v_i$$

Equating these two and using again the linear independence of v_i we obtain $2b = 0$, thus $b = 0$. Thus A is a diagonal matrix with all of its diagonal elements being equal to some number a, therefore it is a scalar multiple of the identity.

Exercise 5
Suppose B is a square matrix with complex entries. Prove that there exists an invertible square matrix A with complex entries such that $A^{-1}BA$ is an upper-triangular matrix.

Solution.
Suppose $u_1, \ldots, u_n$ is the standard basis of $\mathbf{F}^n$, where n is the "dimension" of the matrix B. We know then that there exists a unique $T \in L(\mathbf{F}^n)$ such that $M(T, (u_1, \ldots, u_n)) = B$. Since our field is the complex numbers, there exists a basis $v_1, \ldots, v_n$ of $\mathbf{F}^n$ such that $M(T, (v_1, \ldots, v_n))$ is upper-triangular. We know then that these two matrices of T are related in the following manner:

$$M(T, (v_1, \ldots, v_n)) = A^{-1}M(T, (u_1, \ldots, u_n))A = A^{-1}BA$$

, where $A = M(I, (v_1, \ldots, v_n), (u_1, \ldots, u_n))$. Notice then that for this choice of A, which we know is invertible, $A^{-1}BA$ is indeed upper-triangular.

Exercise 10
Suppose V is an inner product space and $T \in L(V)$. Prove that

$$\text{trace}T^* = \overline{\text{trace}T}$$

Solution.
We can always find an orthonormal basis of V by using the Gram-Schmidt procedure on any basis of it. We also know that if A is the matrix of T with respect to this basis, the matrix of T^* with respect to it is the conjugate transpose of A. We also know that the trace of any operator is equal to the trace of any of its matrices.
Therefore, $\text{trace}T^* = \sum_i \overline{A_{ii}} = \overline{\sum_i A_{ii}} = \overline{\text{trace}T}$

Exercise 14
Suupose $T \in L(V)$ and $c \in \mathbf{F}$. Prove that $\text{trace}(cT) = c\text{trace}(T)$.

Solution.
We know that the trace of any operator is equal to the trace of any of its matrices. Therefore, select any basis of V and let A be the matrix of T with respect to this basis. Then, if we multiply A with c, the result is the matrix that corresponds to the operator cT with respect to this basis. Additionally, it is obvious that since every element of the matrix is multiplied by c, the sum of its diagonal elements is also multiplied by c. Therefore:

$$\text{trace}(cT) = \text{trace}(cA) = c\text{trace}(A) = c\text{trace}T$$

Exercise 15
Suppose $S, T \in L(V)$. Prove that $\text{trace}(ST) = \text{trace}(TS)$.

Solution.
Pick any basis $v_1, \ldots, v_n$ of V and let $A = M(T, (v_1, \ldots, v_n)), B = M(S, (v_1, \ldots, v_n))$. We then have that:

$$\text{trace}(ST) = \text{trace}(BA) = \text{trace}(AB) = \text{trace}(TS)$$

, where we used the fact that for any operator its trace equals the trace of its matrix, the definition of the matrix of the composition of two operators and the proven fact that for two *matrices*, $\text{trace}(BA) = \text{trace}(AB)$.

Exercise 18
Suppose V is an inner product space with orthonormal basis $e_1, \ldots, e_n$ and $T \in L(V)$. Prove that:

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

Conclude that the right side of the equation above is independent of which orthonormal basis $e_1, \ldots, e_n$ is chosen for V.

Solution.
Because $e_1, \ldots, e_n$ is an orthonormal basis, we know that:

$$T^*T(e_i) = \sum_j \langle T^*Te_i, e_j \rangle e_j$$

The diagonal element of the matrix of T^*T with respect to this basis in the i-th column equals the coefficient for e_i when writing $T^*T(e_i)$ as a linear combination of all e_j . Given the above sum, this equals $\langle T^*Te_i, e_i \rangle = \langle Te_i, Te_i \rangle = ||Te_i||^2$. The trace of this matrix (denoted as A) equals the sum of all of these elements, namely:

$$\text{trace}(A) = ||Te_1||^2 + \ldots + ||Te_n||^2$$

Since the trace of an operator equals the trace of any of its matrices, we have that $\text{trace}(T^*T) = ||Te_1||^2 + \ldots + ||Te_n||^2$. Lastly, because the trace is independent of the choice of basis, the right side of this equation must also be independent of which *orthonormal* basis is chosen.

Exercise 19
Suppose V is an inner-product space. Prove that

$$\langle S, T \rangle = \text{trace}(ST^*)$$

defines an inner product on $L(V)$.

Solution.
We need to examine the properties that define an inner product one by one:

- **Positivity**: For any $T \in L(V)$, we have that $\langle T, T \rangle = \text{trace}(TT^*)$. From the previous exercise we observe that this equals $||Te_1||^2 + \ldots + ||Te_n||^2$, where $e_1, \ldots, e_n$ is any orthonormal basis of V. These norms correspond to the inner product on V and are thus nonnegative. Therefore it is also true that $\langle T, T \rangle \geq 0$.

- **Definitiness**: Given a $T \in L(V)$, we need to examine when is it the case that $\langle T, T^* \rangle = 0$. By the observation above, the only way this can happen is if each $||Te_i||$ is zero. By the properties of the inner product on V, this happens iff $Te_i = 0$. Therefore, T must map all vectors of an orthonormal basis to zero, and therefore it must be the zero operator. We conclude that $\langle T, T \rangle = 0$ iff $T = 0$.

- **Homogeneity in the first slot**: For $T \in L(V), a \in \mathbf{F}$ we have that $\langle aT, T \rangle = \text{trace}(aTT^*) = a\text{trace}(TT^*) = a\langle T, T \rangle$, where we used exercise 14.

- **Additivity in the first slot**: For $S, T, R \in L(V)$, we have that:

$$\langle S + T, R \rangle = \text{trace}((S + T)R^*))\text{trace}(SR^* + TR^*) = \text{trace}(SR^*) + \text{trace}(TR^*)$$

$$= \langle S, R \rangle + \langle T, R \rangle$$

  , where we used the additivity of traces and the distributive property of matrix multiplication on addition.

- **Conjugate symmetry**: Given $S, T \in L(V)$ we have that:

$$\langle T, S \rangle = \text{trace}(TS^*) = \text{trace}(S^*T) = \text{trace}(S^*(T^*)^*) =$$

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

  , where we used exercises 10, 15 and the adjoint of the product of two operators.

Therefore, the defined function is indeed an inner product on $L(V)$.

10.B Determinant

Exercise 1
Suppose V is a real vector space. Suppose $T \in L(V)$ has no eigenvalues. Prove that $\det T > 0$.

Solution.
The determinant of T is equal to the product of all the eigenvalues of $T_{\mathbf{C}}$, each raised to the corresponding multiplicity. Since T has no eigenvalues, $T_{\mathbf{C}}$ can have no real eigenvalues. Therefore, all of its eigenvalues are of the form $a + bi, b \neq 0$. We know, however, that if $T_{\mathbf{C}}$ has $a + bi$ as an eigenvalue, it also has $a - bi$, and in fact with the same multiplicity.
Therefore, the product of all of the eigenvalues of $T_{\mathbf{C}}$ can be written as a product $\Pi_j p_j$ of terms of the form $p_j = (a_j + b_j i)(a_j + b_j i) = a_j^2 + b_j^2, b_j \neq 0$. Each of those terms is positive, therefore the product itself is positive, and the determinant of T, which equals this product, is also positive.

Exercise 2
Suppose V is a real vector space with even dimension and $T \in L(V)$. Suppose $\det T < 0$. Prove that T has at least two distinct eigenvalues.

Solution.
Suppose that T has less than two distinct eigenvalues. There are then two cases:

- If T has no eigenvalues, then by the previous exercise $\det T > 0$, which is a contradiction.
- If T has exactly one distinct eigenvalue λ with multiplicity d, then d is either odd or even. If d is even, then the remainining, non-real eigenvalues of $T_{\mathbf{C}}$ have a sum of multiplicities equal to $\dim V - d$ which is also even. Because they come in conjugate pairs of equal multiplicities, the determinant of T must equal $\det T = \lambda^d \Pi_j p_j$, where $p_j = (a_j + b_j i)(a_j + b_j i) = a_j^2 + b_j^2, b_j \neq 0$. The determinant is thus clearly non-negative because d is even and $p_j \geq 0$, therefore we have a contradiction.

  If d is odd, then the remainining, non-real eigenvalues of $T_{\mathbf{C}}$ have a sum of multiplicities equal to $\dim V - d$ which is odd. However, this is a contradiction because the non-real eigenvalues of $T_{\mathbf{C}}$ come in conjugate pairs of equal multiplicities, and thus the sum of those multiplicities cannot be odd.

Since the hypothesis that T has less than two distinct eigenvalues always leads to a contradiction, T must have at least two distinct eigenvalues.

Exercise 4
Suppose $T \in L(V)$ and $c \in \mathbf{F}$. Prove that $\det(cT) = c^{\dim V} \det T$.

Solution.
Suppose first that V is a complex vector space. Suppose $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of T, each one repeated according to its multiplicity. Then $\det T = \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n$. Now let $S = cT$. If $c = 0$, then $S = 0$ and $\det S = 0 = 0^{\dim V} \det T$. Therefore assume that $c \neq 0$. Then, if λ is an eigenvalue of T, there exists a non-zero $v \in V$ such that $T(w) = \lambda w$. Clearly then, $S(v) = cT(v) = c\lambda w$, therefore, $c\lambda$ is an eigenvalue of S with v being a corresponding eigenvector. Conversely, if $S(v) = \mu v$ for a non-zero v then $cT(v) = \mu v \implies T(v) = \frac{\mu}{c} v$, which means that $\frac{\mu}{c}$ is an eigenvalue of T with v being a corresponding eigenvector.
These two observations lead to the conclusion that S has as eigenvalues $c\lambda_1, \dots, c\lambda_n$, where each distinct $c\lambda_j$ is repeated as many times as λ_j was for T (because the corresponding sets of eigenvectors are equal). Clearly then, $\det S = \det(CT) = (c\lambda_1) \cdot (c\lambda_2) \cdot \dots \cdot (c\lambda_n) = c^{\dim V} \det T$, because the sum of all multiplicities equals $\dim V$.
For a real vector space V, it suffices to apply the above argument to $T_{\mathbf{C}}$.

Exercise 6
Suppose A is a block upper-triangular matrix

$$A = \begin{pmatrix} A_1 & \dots & * \\ & \ddots & \vdots \\ 0 & & A_m \end{pmatrix}$$

where each A_j along the diagonal is a square matrix. Prove that

$$\det A = (\det A_1) \cdots (\det A_m)$$

Solution.
Consider the algorithm we use for the computation of the determinant: we compute all possible permutations of the row indices, and then for each of them we take the first-column element of the first row in the permutation, the second-column element of the second column in the permutation etc. and multiply all of them together.
Each of those quantities is multiplied with 1 or -1 and then all of them are added together. In our case here, let $s_1, s_2, \dots s_m$ be the sizes of the square blocks on the diagonal. By the description of the algorithm,

we can see that for a particular permutation term, the first s_1 columns will contribute to the product elements from the first s_1 rows of that permutation.
Observe that if any of those first s_1 permuted row indices is greater than s_1, all of the elements in columns 1 to s_1 are zero. Therefore, the corresponding product will also be zero. This leads to the conclusion that non-zero terms of the sum must the first s_1 rows be a permutation of $(1, 2, \ldots, s_1)$.
The same reasoning applied to any other block results in the conclusion that for permutations $(m_1, m_2, \ldots, m_n)$ whose product is not zero it must be the case that:

$$(m_1, \ldots, m_{s_1}) \in \text{perm}(1, 2, \ldots s_1), (m_{s_1+1}, \ldots, m_{s_1+s_2}) \in \text{perm}(s_1 + 1, \ldots, s_1 + s_2), \ldots,$$

$$(m_{\sum_{k=1}^{m-1} s_i+1}, \ldots, m_n) \in \text{perm}(\sum_{k=1}^{m-1} s_m, \ldots, n)$$

In other words, the permutations can be "decomposed" to sub-permutations of row indices of each block on the diagonal. For permutations with this property, the sign of the permutation must equal the product of the signs of the sub-permutations. This can be proved by induction on the number m of sub-permutations:

- For $m = 2$, we have that all of the indices of the second sub-permutation are larger than all of the indices of the first sub-permutation. Therefore, concatenating them does not create additional wrongly-ordered pairs of indices. This means that the total number of wrong pairs is $k_1 + k_2$, where k_1, k_2 the number of wrong pairs in the two sub-permutations. If both of them are odd or even, i.e. corresponding signs 1, the sum is even, i.e. corresponding sign 1. If exactly one is odd, i.e. one sign -1 and the other 1, the sum is odd, i.e. sign -1. Therefore the sign is always equal to the product of the two sub-permutation signs.
- If this is true for $m \geq 2$ number of sub-permutations, then for $m + 1$ sub-permutations the resulting permutation can be thought of as a permutation resulting from the concatenation of two sub-permutations: the first is the concatenation of m sub-permutations and the second is the last sub-permutation. Applying the induction hypothesis on the first m sub-permutations and then the exact same argument as the induction basis on the concatenation of the last sub-permutation yields the desired result .

Having proved this, we have that:

$$\begin{aligned}
\det A = &\sum_{\substack{(m_1,\ldots,m_{s_1}) \\ \in\text{perm}(1,\ldots,s_1)}} \sum_{\substack{(m_{s_1+1},\ldots,m_{s_1+s_2}) \\ \in\text{perm}(s_1+1,\ldots,s_2)}} \cdots \sum_{\substack{(m_{\sum_{k=1}^{m-1} s_k+1},\ldots,m_n) \\ \in\text{perm}(\sum_{k=1}^{m-1} s_k+1,\ldots,n)}} \text{sign}(m_1, \ldots, m_n) A_{m_1,1} \cdot A_{m_2,2} \cdots A_{m_n,n} \\
= &\sum_{\substack{(m_1,\ldots,m_{s_1}) \\ \in\text{perm}(1,\ldots,s_1)}} \text{sign}(m_1, \ldots, m_{s_1}) A_{m_1,1} \cdots A_{m_{s_1},s_1} \cdot \\
&\sum_{\substack{(m_{s_1+1},\ldots,m_{s_1+s_2}) \\ \in\text{perm}(s_1+1,\ldots,s_2)}} \text{sign}(m_{s_1+1}, \ldots, m_{s_1+s_2}) A_{s_1+1,s_1+1} \cdots A_{m_{s_1+s_2},s_1+s_2} \\
\cdot \ldots &\sum_{\substack{(m_{\sum_{k=1}^{m-1} s_k+1},\ldots,m_n) \\ \in\text{perm}(\sum_{k=1}^{m-1} s_k+1,\ldots,n)}} \text{sign}(m_{\sum_{k=1}^{m-1} s_k+1}, \ldots, m_n) A_{m_{\sum_{k=1}^{m-1} s_k+1}, \sum_{k=1}^{m-1} s_k+1} \cdots A_{m_n,n} \\
= &\det(A_1) \cdots \det(A_m)
\end{aligned}$$

Exercise 8
Suppose V is an inner product space and $T \in L(V)$. Prove that

$$\det T^* = \overline{\det T}$$

Use this to prove that $|\det T| = \det\sqrt{T^*T}$, giving a different proof than was given in 10.47.

Solution.
Consider the characteristic polynomial of $T, p(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0$. We know that the determinant of T equals $a_0(-1)^n$, where $n = \dim V$. Additionally, $p(T) = 0$. Take the adjoint of $p(T)$, which by the properties of the adjoint must be equal to:

$$(p(T))^* = (T^*)^n + \overline{a_{n-1}}(T^*)^{n-1} + \ldots + \overline{a_1}T + \overline{a_0}I$$

Additionally, since the adjoint of the zero map is the zero map, this must equal the zero map as well. This means that for the polynomial $q(z) = z^n + \overline{a_{n-1}}z^{n-1} + \ldots + \overline{a_1}z + \overline{a_0}, q(T^*) = 0$. Therefore, q is the characteristic polynomial of T^*, since it is monic and has degree $\dim V$. Therefore, the determinant of T^* must equal $\det T^* = (-1)^{\dim V}\overline{a_0} = \overline{\det T}$.
Now, consider the following:

$$\det(T^*T) = \det T^* \det T = \overline{\det T}\det T \implies \det(T^*T) = |\det T|^2$$

. However, by definition $\sqrt{T^*T}\sqrt{T^*T} = T^*T$. Therefore:

$$\det(T^*T) = \det(\sqrt{T^*T}\sqrt{T^*T}) \implies \det(\sqrt{T^*T})\det(\sqrt{T^*T}) = |\det T|^2$$

By observing that $\sqrt{T^*T}$ is a positive operator, and thus all of its eigenvalues and its determinant are positive, and by taking square roots we obtain $\det(\sqrt{T^*T}) = |\det T|$.