Homework #4. Solutions to selected problems.

Problems:

0. Let V be a finite-dimensional inner product space over $\mathbb{C}$ and let $A \in \mathcal{L}(V)$ be a normal operator. Prove that A is unitary if and only if all eigenvalues of A have absolute value 1.

Solution: "$\Rightarrow$" Suppose A is unitary, λ an eigenvalue of A and v an eigenvector corresponding to λ. Then

$$\|v\|^2 = \langle v, v\rangle = \langle A^*Av, v\rangle = \langle Av, Av\rangle = \langle \lambda v, \lambda v\rangle = \lambda\overline{\lambda}\|v\|^2 = |\lambda|^2\|v\|^2$$

and hence $|\lambda| = 1$.

"$\Leftarrow$" Conversely, suppose that all eigenvalues of A have absolute value 1. Since unitary operators are normal, by Theorem 7.2 from class there exists an orthonormal basis β for which $[A]_\beta$ is diagonal. The diagonal entries of $[A]_\beta$ are the eigenvalues of A, and since they all have absolute value 1, we must have $[A]_\beta^*[A]_\beta = [A]_\beta[A]_\beta^* = I_n$ (where $n = \dim V$).

On the other hand, since β is orthonormal, $[A]_\beta^* = [A^*]_\beta$. Thus, $[A^*A]_\beta = [A^*]_\beta[A]_\beta = [A]_\beta^*[A]_\beta = I_n$ and similarly $[AA^*]_\beta = I_n$. Hence $A^*A = AA^* = I$ and thus A is unitary.

1. Let V be an inner product space over $\mathbb{C}$ and $A \in GL(V)$, that is, $A \in \mathcal{L}(V)$ is invertible.

(a) Prove that A is unitary if and only if $\langle Ax, Ay\rangle = \langle x, y\rangle$ for all $x, y \in V$.

(b) Now use (a) to prove that A is unitary if and only if $\|Ax\| = \|x\|$ for all $x \in V$.

Solution: (a) "$\Rightarrow$" If A is unitary, then $A^*A = I$, whence $\langle Ax, Ay\rangle = \langle x, A^*Ay\rangle = \langle x, y\rangle$.

"$\Leftarrow$" Suppose now that for all $x, y \in V$ we have $\langle Ax, Ay\rangle = \langle x, y\rangle$. Since A is invertible, we can write every $y \in V$ as $y = A(A^{-1}y)$. Then for all $x, y \in V$ we have

$$\langle Ax, y\rangle = \langle Ax, A(A^{-1}y)\rangle = \langle x, A^{-1}y\rangle.$$

This means that the operator A^{-1} satisfies the definition of the adjoint and thus $A^* = A^{-1}$.

Note: The following argument for "$\Leftarrow$" was presented at the problem session and appeared in the majority of the homework papers. This argument is incomplete in general as it assumes that A^* exists (but is valid for finite-dimensional vector spaces since in that case adjoints always exist):

So assume that $\langle Ax, Ay\rangle = \langle x, y\rangle$, A is invertible and A^* exists. Then as in the proof of $\Rightarrow$ we have $\langle A^*Ax, y\rangle = \langle x, y\rangle$, whence $\langle (A^*Ax - x), y\rangle = 0$. Now fix x and let $z = A^*Ax - x$. We get that $\langle z, y\rangle = 0$ for all $y \in V$. Since the inner product is nondegenerate, we must have $z = 0$. Thus, $A^*Ax = x$ for all $x \in V$, and therefore $A^*A = I$. Since A is invertible, multiplying both sides of this equation by A^{-1} on the right, we get $A^* = A^{-1}$.

Remark: If V is infinite-dimensional, the assumption that A is invertible cannot be eliminated (even if we know that A^* exists). For instance, let $V = \mathbb{C}_0^\infty$ be the vector space of (infinite) finitary complex sequences (where finitary means that only finitely many elements of the sequence are nonzero) with the standard Hermitian dot product and $A : V \to V$ the shift operator given by $A(e_i) = e_{i+1}$ for all $i \geq 1$. It is straightforward to check that A satisfies the equality $\langle Ax, Ay\rangle = \langle x, y\rangle$ for all $x, y \in V$, the adjoint A^* exists and is given by $A^*(e_1) = 0$ and $A^*(e_i) = e_{i-1}$ for $i > 1$. Then $A^*A = I$ (which can be checked directly or deduced from our general argument above), but $AA^* \neq I$ (the operator AA^* fixes e_i for each $i > 1$ but sends e_1 to 0).

2. Let V be an inner product space over $\mathbb{C}$, let $A \in \mathcal{L}(V)$ be unitary, and let $W \subseteq V$ be a *finite-dimensional* subspace of V which is A-invariant (that is, $A(W) \subseteq W$).

(a) Prove that if $A(W) = W$.

(b) Use (a) to prove that $W^\perp$ is also A-invariant.

Solution:(a) Since W is finite-dimensional and $A(W) \subseteq W$, to prove the equality it suffices to show that $\dim A(W) = \dim W$. We apply the rank-nullity theorem the restriction map $A_{|W} : W \to W$. Since $A_{|W}(W) = A(W)$, we get $\dim A(W) = \dim A_{|W}(W) = \dim W - \dim \mathrm{Ker}(A_{|W})$.

Since A is unitary, it is invertible, so we must have $\mathrm{Ker}(A) = 0$ and hence $\mathrm{Ker}(A_{|W}) = 0$. Thus, $\dim A(W) = \dim W$ as desired.

(b) By definition, we need to show that $\langle Az, w\rangle = 0$ for all $w \in W$ and $z \in W^\perp$. Since $A(W) = W$, we can write $w = Av$ for some $v \in W$. But then $\langle Az, w\rangle = \langle Az, Av\rangle = \langle z, v\rangle = 0$ since $z \in W^\perp$ and $v \in W$.

Remark: The assumption that W is finite-dimensional in (a) cannot be eliminated. Indeed, let V be the vector space of finitary complex double-sided sequences (where double-sided means that the sequences are infinite in both directions). This spaces has a natural basis $\{e_i\}_{i\in\mathbb{Z}}$, and again we can turn V into an inner product space using the standard Hermitian dot product. Define the shift operator $A : V \to V$ by $A(e_i) = e_{i+1}$. This time A is unitary. But if we let $W = Span\{e_i : i \geq 1\}$, then $A(W) = Span\{e_i : i \geq 2\}$, so $A(W) \subseteq W$ while $A(W) \neq W$.

Note that in the proof of (b) we did not directly use the fact that W is finite-dimensional, only that it satisfies $A(W) = W$.

4. Let V be a finite-dimensional inner product space over $\mathbb{C}$ and let H and G be Hermitian forms on V.

(a) Assume that G is positive-definite. Prove that there exists a basis β of V such that $[H]_\beta$ and $[G]_\beta$ are both diagonal (equivalently, if $A, B \in Mat_n(\mathbb{C})$ are Hermitian matrices and A is positive definite, there exists $P \in GL_n(\mathbb{C})$ such that P^*AP and P^*BP are both diagonal).

Solution: Since G is Hermitian positive-definite, we can define an inner product structure on V by setting $\langle x, y\rangle = G(x, y)$. By Corollary 8.6, there exists an orthonormal basis β (with respect to this inner product) such that $[H]_\beta$ is diagonal. On the other hand, by definition of the orthonormal basis we must have $[G]_\beta = I_n$, where $n = \dim V$. Thus, $[G]_\beta$ and $[H]_\beta$ are both diagonal.

5.

(a) Let $A \in Mat_n(\mathbb{C})$ be arbitrary. Prove that there exists a unitary matrix U such that $U^{-1}AU = U^*AU$ is upper-triangular.

(b) Use (a) to give a short proof of Theorem 8.5 from class in the case when A is Hermitian or unitary (without referring to Theorem 7.2)

Solution: (a) Contrary to the hint, we will phrase the solution in terms of matrices. We argue by induction on $n = \dim(V)$. The base case $n = 1$ is trivial. Fix $n > 1$, and assume that the assertion holds for matrices of smaller size.

Let λ be any eigenvalue of A and v a corresponding eigenvector. After rescaling, we can assume that $\|v\| = 1$. Let β be any orthonormal basis of V whose first element is v (to get such a basis simply take any orthonormal basis for $(\mathbb{C}v)^\perp$ and add v at the beginning). Let U_1 be the matrix whose columns are elements of β. Then U_1 is unitary since β is orthonormal. Also, the matrix $U_1^{-1}AU_1$ has the block form $\begin{pmatrix} \lambda & y \\ 0 & B \end{pmatrix}$ where 0 is a column of 0's of height $n-1$, y is some row of length $n-1$ and $B \in Mat_{n-1}(\mathbb{C})$. Indeed, $U_1^{-1}AU_1 = [\widetilde{A}]_\beta$ where $\widetilde{A}$ is the operator represented by A in the standard basis, and the assertion follows from the fact that the first element v of β satisfies $\widetilde{A}v = \lambda v$.

Now by induction hypothesis there exists a unitary $(n-1)\times(n-1)$ matrix V such that $V^{-1}BV$ is upper-triangular. Now consider the block-diagonal $n\times n$-matrix $U_2 = \begin{pmatrix} 1 & 0 \\ 0 & V \end{pmatrix}$. Since V is unitary, it is straightforward to check that U_2 is unitary as well.

Finally, let $U = U_1U_2$. Since unitary matrices form a group (which follows easily from the definition), U is also unitary. Also

$$U^{-1}AU = U_2^{-1}(U_1^{-1}AU_1)U_2 = \begin{pmatrix} 1 & 0 \\ 0 & V \end{pmatrix}^{-1} \begin{pmatrix} \lambda & y \\ 0 & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & V \end{pmatrix}$$

Using the formula for multiplying block matrices, we get that the above product is equal to $\begin{pmatrix} \lambda & z \\ 0 & V^{-1}BV \end{pmatrix}$ for some row vector z of length $n-1$. Since $V^{-1}BV$ is upper-triangular, the matrix $\begin{pmatrix} \lambda & z \\ 0 & V^{-1}BV \end{pmatrix}$ is upper-triangular as well, which finishes the proof.

(b) First we claim that if A is Hermitian (respectively unitary) and U is unitary, then U^*AU is Hermitian (respectively unitary). Indeed, if $A^* = A$, then $(U^*AU)^* = U^*A^*U = U^*AU$ (here we did not use that U is unitary) and if A is unitary, then $U^*AU = U^{-1}AU$ is unitary since unitary matrices form a group.

Thus, by (a) we just need to show that an upper-triangular matrix which is either Hermitian or unitary must be diagonal. If A is upper-triangular, then A^* is lower-triangular, so $A^* = A$ forces A to be diagonal.

Also invertible upper-triangular $n\times n$ matrices form a subgroup of $GL_n(\mathbb{C})$, so if A is upper-triangular and unitary (hence invertible), A^{-1} must be upper-triangular. Hence again $A^* = A^{-1}$ forces A to be diagonal.