MATH 325 Q1: LINEAR ALGEBRA III, PART 9

Hermitian and unitary linear maps

Notation. Throughout this section all matrices etc. are viewed over the complex numbers $\mathbb{C}$. As usual, we denote by $\bar{z}$ the complex conjugate of $z \in \mathbb{C}$. We will be considering $\mathbb{C}^n$ as the inner product space with the usual Euclidean scalar product $\langle -, - \rangle$. We let A^T be the transpose and by $\bar{A}$ the complex conjugate of a matrix $A = (a_{ij})$, *i.e.* the ij-entry of A^T is a_{ji} and of $\bar{A}$ is $\bar{a}_{ij}$.

Note that with this notation we have

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}^T = (x_1 \ldots x_n)$$

and

$$\overline{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}} = \begin{pmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_n \end{pmatrix}$$

for a vector $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{C}^n$. In particular we have then

$$\langle x, y \rangle = \sum x_i \bar{y}_i = x^T \cdot \bar{y}$$

for all $x, y \in \mathbb{C}^n$.

The adjoint linear map

Let $(V, \langle -, - \rangle)$ be an inner product space, and let $\alpha : V \to V$ be a $\mathbb{C}$-linear map. Then for a fixed vector $w \in V$ the map

$$f_w : V \longrightarrow \mathbb{C}, \; v \longmapsto \langle \alpha(v), w \rangle$$

is $\mathbb{C}$-linear (because α is $\mathbb{C}$-linear and because of **(L1)**). Therefore there exists a unique vector $y = y(w) \in V$ (depending on w) such that

$$f_w(v) = \langle v, y \rangle .$$

We define $\alpha^*(w) := y = y(w)$ for $w \in V$. This defines a map

$$\alpha^* : V \longrightarrow V, \; w \longmapsto \alpha^*(w),$$

such that

$$\langle \alpha(v), w \rangle = \langle v, \alpha^*(w) \rangle$$

for all $v, w \in V$.

In two lemmas below we summarize main properties of α^*.

Lemma. *The map α^* is $\mathbb{C}$-linear.*

Proof. Let $w_1, w_2 \in V$ and $\lambda_1, \lambda_2 \in \mathbb{C}$. Then we have:

$$\begin{aligned} f_{\lambda_1 w_1 + \lambda_2 w_2}(v) &= \langle \alpha(v), \lambda_1 w_1 + \lambda_2 w_2 \rangle \\ &= \bar{\lambda}_1 \cdot \langle \alpha(v), w_1 \rangle + \bar{\lambda}_2 \cdot \langle \alpha(v), w_2 \rangle && \text{by } \mathbf{(L2)} \\ &= \bar{\lambda}_1 \cdot \langle v, \alpha^*(w_1) \rangle + \bar{\lambda}_2 \cdot \langle v, \alpha^*(w_2) \rangle && \text{by definition of } \alpha^* \\ &= \langle v, \lambda_1 \cdot \alpha^*(w_1) + \lambda_2 \cdot \alpha^*(w_2) \rangle && \text{by } \mathbf{(L2).} \end{aligned}$$

On the other hand by the definition of the map α^* we have also

$$f_{\lambda_1 w_1 + \lambda_2 w_2}(v) = \langle \alpha(v), \lambda_1 w_1 + \lambda_2 w_2 \rangle = \langle v, \alpha^*(\lambda_1 w_1 + \lambda_2 w_2) \rangle,$$

and so

$$\langle v, \alpha^*(\lambda_1 w_1 + \lambda_2 w_2) \rangle = \langle v, \lambda_1 \cdot \alpha^*(w_1) + \lambda_2 \cdot \alpha^*(w_2) \rangle .$$

This implies that

$$\alpha^*(\lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 \cdot \alpha^*(w_1) + \lambda_2 \cdot \alpha^*(w_2),$$

as required. □

Definition. *The linear map α^* is is called the adjoint of α (with respect to the inner product $\langle -, - \rangle$).*

We collect some more elementary properties of the adjoint in the following lemma.

Lemma. *Let $(V, \langle - \rangle)$ be an inner product space. Then:*

(i) $(\mathrm{id}_V)^* = \mathrm{id}_V$.

(ii) $(\alpha^*)^* = \alpha$ *for all linear maps* $\alpha : V \longrightarrow V$.

(iii)

$$(\lambda_1 \cdot \alpha_1 + \lambda_2 \cdot \alpha_2)^* = \bar{\lambda}_1 \cdot \alpha_1^* + \bar{\lambda}_2 \cdot \alpha_2^*$$

for all linear maps $\alpha_1, \alpha_2 : V \longrightarrow V$ *and all scalars* $\lambda_1, \lambda_2 \in \mathbb{C}$.

(iv) $(\beta \circ \alpha)^* = \alpha^* \circ \beta^*$ *for all linear maps* $\alpha, \beta : V \longrightarrow V$.

The assertions (i) and (iv) imply in particular that if $\alpha : V \longrightarrow V$ is an isomorphism then also α^* is an isomorphism and

$$(\alpha^*)^{-1} = (\alpha^{-1})^* .$$

Proof. (i) follows directly form the definition.

For (ii): we will repeatedly use below the following property:

$\mathcal{P}$: *if w is a vector such that $\langle v, w\rangle = 0$ for all vectors v then $w = 0$; or equivalently if w_1, w_2 are vectors such that $\langle v, w_1\rangle = \langle v, w_2\rangle$ for all vectors v then $w_1 = w_2$.*

We have using axiom **(H)** and the definition of α^*:

$$\langle \alpha^*(v), w\rangle \; = \; \overline{\langle w, \alpha^*(v)\rangle} \; = \; \overline{\langle \alpha(w), v\rangle} \; = \; \langle v, \alpha(w)\rangle$$

for all $v, w \in V$. On the other hand we also have

$$\langle \alpha^*(v), w\rangle = \langle v, (\alpha^*)^*(w)\rangle$$

and so by property $\mathcal{P}$ we have

$$\alpha(w) = (\alpha^*)^*(w)$$

for all vectors w and this implies $\alpha = (\alpha^*)^*$.

For (iii) we have

$$\begin{aligned}\langle \lambda_1 \cdot \alpha_1(v) + \lambda_2 \cdot \alpha_2(v), w\rangle &= \lambda_1 \cdot \langle \alpha_1(v), w\rangle + \lambda_2 \cdot \langle \alpha_2(v), w\rangle \quad &\text{by } \textbf{(L1)}\\ &= \lambda_1 \cdot \langle v, \alpha_1^*(w)\rangle + \lambda_2 \cdot \langle v, \alpha_2^*(w)\rangle \quad &\text{def. adj.}\\ &= \langle v, \bar{\lambda}_1 \cdot \alpha_1^*(w) + \bar{\lambda}_2 \cdot \alpha_2^*(w)\rangle \quad &\text{by } \textbf{(L2)}\end{aligned}$$

for all $v, w \in V$. On the other hand side we have by definition that

$$\langle \lambda_1 \cdot \alpha_1(v) + \lambda_2 \cdot \alpha_2(v), w\rangle = \langle v, (\lambda_1 \cdot \alpha_1 + \lambda_2 \cdot \alpha_2)^*(w)\rangle$$

The property $\mathcal{P}$ then implies

$$\bar{\lambda}_1 \cdot \alpha_2^* + \bar{\lambda}_2 \cdot \alpha_2^* = (\lambda_1 \cdot \alpha_1 + \lambda_2 \cdot \alpha_2)^*,$$

as required.

Finally we prove (iv). Using the definition of the adjoint map we compute:

$$\langle v, (\beta \circ \alpha)^*(w)\rangle \; = \; \langle \beta(\alpha(v)), w\rangle \; = \; \langle \alpha(v), \beta^*(w)\rangle \; = \; \langle v, \alpha^*(\beta^*(w))\rangle\,.$$

As above this implies

$$(\beta \circ \alpha)^* = \alpha^* \circ \beta^*.$$

For the last assertion, we if $\alpha : V \longrightarrow V$ is an isomorphism then there exists α^{-1}, such that $\alpha^{-1} \circ \alpha = \mathrm{id}_V$, and so by (i) and (iv) we have

$$\mathrm{id}_V \; = \; \mathrm{id}_V^* \; = \; (\alpha^{-1} \circ \alpha)^* \; = \; \alpha^* \circ (\alpha^{-1})^*\,.$$

But this implies since V is finite dimensional that also α^* is invertible and the inverse is $(\alpha^{-1})^*$. □

Example. Let $A = (a_{ij})$ be a complex $n \times n$-matrix, and

$$\alpha_A : \mathbb{C}^n \longrightarrow \mathbb{C}^n, \quad v \mapsto A \cdot v$$

be the linear map defined by A. We want to compute the matrix of the adjoint linear map α_A^* with respect to the standard basis. Recall that we consider $\mathbb{C}^n$ as the inner Euclidean product space with the usual scalar product.

Let $x, y \in \mathbb{C}$. Then

$$\langle \alpha_A(x), y \rangle = \langle Ax, y \rangle = (Ax)^T \cdot \bar{y} = x^T \cdot A^T \cdot \bar{y} = x^T \cdot (\overline{\bar{A}^T \cdot y}) = x^T \cdot \overline{A^* y} = \langle x, A^* y \rangle ,$$

where we have set $A^* := \bar{A}^T$. But also, we have

$$\langle \alpha_A(x), y \rangle = \langle x, \alpha^*(y) \rangle .$$

It follows by property $\mathcal{P}$ that $\alpha^*(y) = A^* y$ for all vectors y, hence the matrix of α_A^* with respect to the standard basis is $A^* = \bar{A}^T$.

Definition. *The matrix $A^* = \bar{A}^T$ is called the adjoint matrix of A.*

Example. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then

$$A^* = \bar{A}^T = \begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix} .$$

Lemma. (i) *One has $(A^*)^* = A$.*
(ii) *Let $A, B \in M_n(\mathbb{C})$. Then $(AB)^* = B^* A^*$.*

Proof. This is obvious. □

Definition. *Let $(V, \langle -, \rangle)$ be an inner product space and $\alpha : V \longrightarrow V$ be a $\mathbb{C}$-linear map. Then:*

(i) *α is a called hermitian or self-adjoint if $\alpha = \alpha^*$.*
(ii) *α is a called unitary if $\alpha^* \circ \alpha = \mathrm{id}_V$. Note that then one also has $\alpha \circ \alpha^* = \mathrm{id}_V$.*

Definition. *We call a $n \times n$-matrix A hermitian (respectively unitary) if α_A is hermitian (respectively unitary), where we consider $\mathbb{C}^n$ as an inner product space with the usual Euclidean inner product.*

Remark. Since A^* is the matrix of α_A^* we see that A is hermitian if and only if $A = A^*$, and unitary if and only if $A \cdot A^* = A^* \cdot A = \mathrm{I}_n$.

Lemma. *Let $(V, \langle -, - \rangle)$ be an inner product space and $\alpha : V \longrightarrow V$ a linear map. Let further $W \subseteq V$ an α-invariant subspace, i.e. $\alpha(w) \in W$ for all $w \in W$. If α is hermitian then $\alpha|_W$ is also hermitian.*

Proof. The subspace W with the restriction of the inner product $\langle -, - \rangle$ is also an inner product space. The adjoint of the linear map

$$\alpha|_W : W \longrightarrow W, \; w \longmapsto \alpha(w)$$

is the unique map $\beta : W \longrightarrow W$, such that

$$\langle \alpha(v), w \rangle = \langle v, \beta(w) \rangle$$

for all $v, w \in W$. If α is hermitian then $\alpha^*(w) = \alpha(w) \in W$ for all $w \in W$, and so $\beta = \alpha|_W$ has this property. Consequently, $\big(\alpha|_W\big)^* = \alpha|_W$, *i.e.* $\alpha|_W$ is hermitian. $\square$

Unitary matrices and orthonormal bases

Let $u_1, \ldots, u_n$ be an orthonormal basis of $\mathbb{C}^n$ (not necessary standard) with respect to the usual scalar product. By definition we have

$$u_i^T \cdot \bar{u}_j = \langle u_i, u_j \rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \,. \end{cases} \tag{1}$$

Let $u_i = \begin{pmatrix} u_{1i} \\ u_{2i} \\ \vdots \\ u_{ni} \end{pmatrix}$ and $U := (u_{ij})$, *i.e.* the ith column of U is u_i. Thus U is the base change matrix from the basis $u_1, \ldots, u_n$ to the standard one.

Lemma. *The matrix U is unitary.*

Proof. Using (1) we conclude that $U^T \cdot \bar{U} = \mathrm{I}_n$, and so since $\bar{\mathrm{I}}_n = \mathrm{I}_n$ we get

$$\mathrm{I}_n = \bar{\mathrm{I}}_n = \overline{U^T \cdot \bar{U}} = \bar{U}^T \cdot U = U^* \cdot U \,,$$

i.e. U is an unitary matrix. $\square$

Remark. *Going this reasoning backwards we see that if U is unitary then the columns of U are an orthonormal basis of $\mathbb{C}^n$.*

Remark. *Using that $U^* = \bar{U}^T$ one can easily see that the rows of an unitary matrix are an orthonormal basis of $\mathbb{C}^n$ and vice versa.*

Theorem. *Let $(V, \langle -, \rangle)$ be an inner product space and $\alpha : V \longrightarrow V$ a $\mathbb{C}$-linear map. Then the following assertions are equivalent:*

(i) $\|\alpha(v)\| = \|v\|$ *for all* $v \in V$.

(ii) $\langle \alpha(v), \alpha(w) \rangle = \langle v, w \rangle$ *for all* $v, v \in V$.

(iii) α *is unitary.*

(iv) *If* $v_1, \dots, v_n$ *is an orthonormal basis of* V *then also* $\alpha(v_1), \dots, \alpha(v_n)$ *is an orthonormal basis.*

This implies in particular, that if α *is unitary and* $W \subseteq V$ *is a* α*-invariant subspace then also the restriction*

$$\alpha|_W : W \longrightarrow W, \; w \longmapsto \alpha(w)$$

is unitary.

Proof. `(i) implies (ii)`: We have by (i) that

$$\langle v+w, v+w\rangle = \langle \alpha(v+w), \alpha(v+w)\rangle$$

and so since also $\|\alpha(v)\| = \|v\|$ and $\|\alpha(w)\| = \|w\|$ we get by **(L1)**, **(L2)** and **(H)**:

$$\langle \alpha(v), \alpha(w)\rangle + \overline{\langle \alpha(v), \alpha(w)\rangle} = \langle v, w\rangle + \overline{\langle v, w\rangle},$$

i.e. the real part of $\langle \alpha(v), \alpha(w)\rangle$ and $\langle v, w\rangle$ coincide for all $v, w \in V$. Setting $i := \sqrt{-1} \in \mathbb{C}$ we get from

$$\langle v + i\cdot w, v + i\cdot w\rangle = \langle \alpha(v + i\cdot w), \alpha(v + i\cdot w)\rangle$$

by the same reasoning and using the equality $\|i\cdot v\| = |i|\cdot\|v\| = \|v\|$, we see that

$$-i\langle \alpha(v), \alpha(w)\rangle + i\overline{\langle \alpha(v), \alpha(w)\rangle} = -i\langle v, w\rangle + i\overline{\langle v, w\rangle},$$

and so that also the imaginary parts of $\langle \alpha(v), \alpha(w)\rangle$ and $\langle v, w\rangle$ coincide. Therefore we have proven

$$\langle \alpha(v), \alpha(w)\rangle = \langle v, w\rangle$$

for all $v, w \in V$.

`(ii) implies (iii)`: We have $\langle v, w\rangle = \langle \alpha(v), \alpha(w)\rangle$ by (ii) and therefore

$$\langle v, w\rangle = \langle \alpha(v), \alpha(w)\rangle = \langle v, \alpha^*(\alpha(w))\rangle$$

for all $v, w \in V$. Using the property $\mathcal{P}$ we conclude that $\alpha^* \circ \alpha = \mathrm{Id}_V$.

`(iii) implies (iv)`: That α is unitary means in particular that α is an isomorphism and so $\alpha(v_1), \dots, \alpha(v_n)$ is also a basis of V if $v_1, \dots, v_n$ is one. Hence we are left to show that $\alpha(v_1), \dots, \alpha(v_n)$ are pairwise orthogonal and have length 1.

But by the definition of the adjoint map we have

$$\langle \alpha(v_i), \alpha(v_j)\rangle = \langle v_i, \alpha^*(\alpha(v_j))\rangle = \langle v_i, v_j\rangle$$

(the latter since $\alpha^* \circ \alpha = \mathrm{id}_V$), and so $\langle \alpha(v_i), \alpha(v_j)\rangle = 1$ if $i = j$ and is 0 if $i \neq j$ since $v_1, \dots, v_n$ is an orthonormal basis.

`(iv) implies (i)`: Let $v_1 = \frac{v}{\|v\|}$. Its length is equal to 1. By the Gram-Schmidt process we can complete it to an orthonormal basis $\{v_1, \ldots, v_n\}$ of V. Then by (iv) we have

$$\langle \alpha\left(\frac{v}{\|v\|}\right), \alpha\left(\frac{v}{\|v\|}\right)\rangle = \langle \alpha(v_1), \alpha(v_1)\rangle = 1$$

implying

$$\langle \alpha(v), \alpha(v)\rangle = \|v\|^2 = \langle v, v\rangle,$$

as required.

The last assertion of the theorem is a consequence of (ii) $\iff$ (iii): The map α is unitary if and only if $\langle v, w\rangle = \langle \alpha(v), \alpha(w)\rangle$ for all $v, w \in V$, and so in particular for all v, w in the subspace W. Hence also the restriction $\alpha|_W$ is unitary. □

Corollary. *Let $(V, \langle -, -\rangle)$ be an inner product space and $\alpha : V \to V$ a unitary $\mathbb{C}$-linear map. Then for all eigenvalues λ of α one has $|\lambda| = 1$.*

Proof. Let v be an eigenvector for the eigenvalue λ. Then we have by (i) in the theorem above

$$\|v\| = \|\alpha(v)\| = \|\lambda \cdot v\| = |\lambda| \cdot \|v\|,$$

and so $|\lambda| = 1$ since $\|v\| \neq 0$. □

For hermitian maps we have the following.

Lemma. *Let $(V, \langle -, -\rangle)$ be an inner product space and $\alpha : V \longrightarrow V$ a hermitian $\mathbb{C}$-linear map. Then all eigenvalues of α are real.*

Proof. Let λ be an eigenvalue of α with corresponding eigenvector $0 \neq v \in V$. We want to show that $\lambda \in \mathbb{R}$. For this we observe first that

$$\alpha(\alpha(v)) = \alpha(\lambda v) = \lambda^2 \cdot v$$

and so v is an eigenvectors of $\alpha \circ \alpha$ for the eigenvalue λ^2. Using this we compute:

$$\begin{aligned}
(\lambda \cdot \bar{\lambda}) \cdot \langle v, v\rangle &= \langle \lambda v, \lambda v\rangle && \text{by } \textbf{(L1)} \text{ and } \textbf{(L2)} \\
&= \langle \alpha(v), \alpha(v)\rangle && \text{since } v \text{ is eigenvector for } \lambda \\
&= \langle v, \alpha^*(\alpha(v))\rangle && \text{since } \alpha^* \text{ is the adjoint of } \alpha \\
&= \langle v, \alpha(\alpha(v))\rangle && \text{since } \alpha^* = \alpha \\
&= \langle v, \lambda^2 \cdot v\rangle && \text{by the remark above} \\
&= \bar{\lambda}^2 \cdot \langle v, v\rangle && \text{by } \textbf{(L2)}.
\end{aligned}$$

Since $v \neq 0$ we have $\langle v, v \rangle > 0$ by **(P)** and so the above equation implies $\lambda \cdot \bar{\lambda} = \bar{\lambda}^2$ and therefore $\lambda = \bar{\lambda}$, *i.e.* $\lambda \in \mathbb{R}$. $\square$