7. The Dual and Adjoint of a Linear Map

7.1 Definition: For two vector spaces U and V over a field $\mathbb{F}$, we write $\mathrm{Hom}(U,V)$ for the vector space of linear maps $L:U\to V$. For a vector space U over a field $\mathbb{F}$, the **dual** of U is the vector space
$$U^*=\mathrm{Hom}(U,\mathbb{F}).$$

7.2 Theorem: (*Dual Basis*) *Let U be an n-dimensional vector space over a field $\mathbb{F}$. Let $\mathcal{A}=\{u_1,\cdots,u_n\}$ be a basis for U. For each index k, let $f_k\in U^*$ be the linear map $f_k:U\to\mathbb{F}$ such that $f_k(u_i)=\delta_{k,i}$. Then the set $\mathcal{F}=\{f_1,\cdots,f_n\}$ is a basis for U^*. Also, for $x\in U$ and $g\in U^*$ we have*
$$[x]_{\mathcal{A}}=\begin{pmatrix}f_1(x)\\ \vdots\\ f_n(x)\end{pmatrix}\quad\text{and}\quad[g]_{\mathcal{F}}=\begin{pmatrix}g(u_1)\\ \vdots\\ g(u_n)\end{pmatrix}.$$

Proof: For $x=\sum_{i=1}^n t_iu_i\in U$ we have
$$f_k(x)=f_k\big(\sum_{i=1}^n t_iu_i\big)=\sum_{i=1}^n t_i\,f_k(u_i)=\sum_{i=1}^n t_i\delta_{k,i}=t_k$$
and so $[x]_{\mathcal{A}}=\big(f_1(x),\cdots,f_n(x)\big)^T$. For $g=\sum_{i=1}^n t_if_i\in\mathrm{Span}\,\mathcal{F}$, we have
$$g(u_k)=\Big(\sum_{i=1}^n t_if_i\Big)(u_k)=\sum_{i=1}^n t_if_i(u_k)=\sum_{i=1}^n f_i\delta_{i,k}=t_k.$$
It follows that $\mathcal{F}$ is linearly independent because if $\sum_{i=1}^n t_if_i=0$ then $t_k=\Big(\sum_{i=1}^n t_if_i\Big)(u_k)=0$ for all k, and it follows that $\mathcal{F}$ spans U^* because given any $g\in U^*$ we can let $t_k=g(u_k)$ and then we have $g(u_k)=\Big(\sum_{i=1}^n g(u_i)f_i\Big)(u_k)$ for all k, and this implies that $g=\sum_{i=1}^n g(u_i)f_i$ so that $g\in\mathrm{Span}\,\mathcal{F}$. It also follows that $[g]_{\mathcal{F}}=\big(g(u_1),\cdots,g(u_n)\big)^T$.

7.3 Definition: The basis $\mathcal{F}$ in the above theorem is called the **dual basis** of $\mathcal{A}$ for U^*.

7.4 Remark: If U is a countable dimensional vector space over $\mathbb{F}$ and $\mathcal{A}=\{u_1,u_2,\cdots\}$ is a basis for U, then for each index k we can still let $f_k\in U^*$ be the linear map $f_k:U\to\mathbb{F}$ given by $f_k(u_i)=\delta_{k,i}$. Then the set $\mathcal{F}=\{f_1,f_2,\cdots\}$ is still linearly independent, but it no longer spans U^*. In this case we have
$$\mathrm{Span}\,\mathcal{F}\cong\mathbb{F}^\infty\quad\text{and}\quad\mathcal{U}^*\cong\mathbb{F}^\omega.$$
Indeed every $g\in U^*$ is uniquely determined by the values $g(u_i)$, and we can define a vector space isomorphism $\phi_{\mathcal{A}}:U^*\to\mathbb{F}^\omega$ by $\phi_{\mathcal{A}}(g)=\big(g(u_1),g(u_2),\cdots\big)$.

More generally, if U is any vector space over $\mathbb{F}$ and $\mathcal{A}$ is a basis, then for each $u\in\mathcal{A}$ we can let $f_u\in U^*$ be the unique linear map $f_u:U\to\mathbb{F}$ such that $f_u(u)=1$ and $f_u(v)=0$ for $v\in\mathcal{A}$ with $v\neq u$. Then the set $\mathcal{F}=\big\{f_u\big|u\in\mathcal{A}\big\}$ is linearly independent, but when U is infinite dimensional we have $\mathrm{Span}\,\mathcal{F}\subsetneqq U^*$.

7.5 Theorem: *(Double Dual) Let U be a vector space over a field $\mathbb{F}$. Define $\phi : U \to (U^*)^*$ by $\phi(u)(g) = g(u)$ for $u \in U$ and $g \in U^*$. Then*

(1) ϕ is an injective linear map, and

(2) if U is finite dimensional then ϕ is bijective.

Proof: The map ϕ is linear because for all $u, v \in U$ we have

$$\phi(u+v)(g) = g(u+v) = g(u) + g(v) = \phi(u)(g) + \phi(v)(g) = \big(\phi(u) + \phi(v)\big)(g)$$

for all $g \in U^*$ so that $\phi(u+v) = \phi(u) + \phi(v)$, and because for all $u \in U$ and all $t \in \mathbb{F}$ we have

$$\phi(tu)(g) = g(tu) = tg(u) = t\big(\phi(u)(g)\big) = \big(t\phi(u)\big)(g)$$

for all $g \in U^*$ so that $\phi(tu) = t\phi(u)$. The map ϕ is injective because, for $u \in U$, if $\phi(u) = 0$ then $\phi(u)(g) = 0$ for all $g \in U^*$, and hence $g(u) = 0$ for all $g \in U^*$, and this implies that $u = 0$ (since if $u \neq 0$ we can construct $g \in U^*$ such that $g(u) \neq 0$ as follows: extend $\{u\}$ to a basis $\mathcal{A}$ for U, then define $g \in U^*$ to be the linear map $g : U \to \mathbb{F}$ given by $g(u) = 1$ and $g(v) = 0$ for $v \in \mathcal{A}$ with $v \neq u$). This proves Part (1).

Suppose that U is finite dimensional. By the Dual Basis Theorem, we know that $\dim U = \dim U^*$ and $\dim U^* = \dim(U^*)^*$. Since $\phi : U \to (U^*)^*$ is injective and $\dim U = \dim(U^*)^*$, it follows that ϕ is bijective. This proves Part (2).

7.6 Definition: The map $\phi : U \to (U^*)^*$ of the above theorem, given by $\phi(u)(g)$ is called the **evaluation map**.

7.7 Definition: Let U and V be vector spaces over a field $\mathbb{F}$. Let $L : U \to V$ be a linear map. The **dual** of the map L is the linear map $L^T : V^* \to U^*$ given by $L^T(g) = g \circ L$ so that $L^T(g)(u) = g\big(L(u)\big)$ for all $g \in V^*$ and $u \in U$.

7.8 Theorem: *Let U and V be finite dimensional vector spaces over a filed $\mathbb{F}$. Let $\mathcal{A}$ and $\mathcal{B}$ be ordered bases for U and V. Let $\mathcal{F}$ and $\mathcal{G}$ be the dual bases for U^* and V^*. Let $L : U \to V$ be a linear map. Then*

$$\big[L^T\big]_{\mathcal{F}}^{\mathcal{G}} = \big([L]_{\mathcal{B}}^{\mathcal{A}}\big)^T.$$

Proof: Let $\mathcal{A} = \{u_1, \cdots, u_k\}$, $\mathcal{B} = \{v_1, \cdots, v_l\}$, $\mathcal{F} = \{f_1, \cdots, f_k\}$ and $\mathcal{G} = \{g_1, \cdots, g_l\}$. Then using the formulas for coefficient vectors from the Dual Basis Theorem, we have

$$[L]_{\mathcal{B}}^{\mathcal{A}} = \Big(\big[L(u_1)\big]_{\mathcal{B}}, \cdots, \big[L(u_l)\big]_{\mathcal{B}}\Big) = \begin{pmatrix} g_1(Lu_1) & g_1(Lu_2) & \cdots & g_1(Lu_k) \\ g_2(Lu_1) & g_2(Lu_2) & \cdots & g_2(Lu_k) \\ \vdots & \vdots & & \vdots \\ g_l(Lu_1) & g_l(Lu_2) & \cdots & g_l(Lu_k) \end{pmatrix}$$

and

$$\big[L^T\big]_{\mathcal{F}}^{\mathcal{G}} = \Big(\big[L^T(g_1)\big]_{\mathcal{F}}, \cdots, \big[L^T(g_l)\big]_{\mathcal{F}}\Big) = \begin{pmatrix} g_1(Lu_1) & g_2(Lu_1) & \cdots & g_l(Lu_1) \\ g_1(Lu_2) & g_2(Lu_2) & \cdots & g_l(Lu_2) \\ \vdots & \vdots & & \vdots \\ g_1(Lu_k) & g_2(Lu_k) & \cdots & g_l(Lu_k) \end{pmatrix}.$$

7.9 Definition: Let W be a vector space over a field $\mathbb{F}$. For a subspace $U \subseteq W$, the **annihilator** of U in W^* is the space

$$U^\circ = \left\{ g \in W^* \,\middle|\, g(x) = 0 \text{ for all } x \in U \right\}.$$

7.10 Theorem: *Let W be a finite dimensional vector space over a field $\mathbb{F}$. Let $U \subseteq W$ be a subspace. Then*

$$\dim U + \dim U^\circ = \dim W.$$

Proof: Let $\{u_1, u_2 \cdots, u_k\}$ be an ordered basis for U. Extend this to an ordered basis $\{u_1, \cdots, u_k, v_1, \cdots, v_l\}$ for W. Let $\{f_1, \cdots, f_k, g_1, \cdots, g_l\}$ be the dual basis for W^*. We claim that $\{g_1, \cdots, g_l\}$ is a basis for U°. Since $g_j(u_i) = 0$ for all $1 \leq i \leq k$, we see that each $g_j \in U^\circ$ so we have $\text{Span}\,\{g_1, \cdots, g_l\} \subseteq U^\circ$. For $h \in U^\circ$, say $h = \sum_{i=1}^{k} s_i f_i + \sum_{i=1}^{l} t_i g_i$, we have $s_j = h(u_j) = 0$ for all indices j so that $h = \sum_{i=1}^{l} t_i g_i \in \text{Span}\,\{g_1, \cdots, g_l\}$. Thus $\text{Span}\,\{g_1, \cdots, g_l\} = U^\circ$, and so $\{g_1, \cdots, g_l\}$ is a basis for U°, as claimed.

7.11 Theorem: *Let U be a finite dimensional inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Define $\phi_U : U \to U^*$ by $\phi(u)(x) = \langle x, u \rangle$ for $u, x \in U$. Then*
(1) if $\mathbb{F} = \mathbb{R}$ then ϕ_U is a vector space isomorphism, and
(2) if $\mathbb{F} = \mathbb{C}$ then ϕ_U is conjugate-linear and bijective.

Proof: Let $\phi = \phi_U$. The map ϕ is well-defined because for $u \in U$, the map $\phi(u) : U \to F$ given by $\phi(u)(x)$ is linear in x so that $\phi(u) \in U^*$. The map ϕ is linear when $\mathbb{F} = \mathbb{R}$ and conjugate-linear when $\mathbb{F} = \mathbb{C}$ because for $u, v \in U$ and $t \in \mathbb{F}$ we have

$$\begin{aligned} \phi(u+v)(x) &= \langle x, u+v \rangle = \langle x, u \rangle + \langle x, v \rangle = \phi(u)(x) + \phi(v)(x) \text{ , and} \\ \phi(tu)(x) &= \langle x, tu \rangle = t\langle x, u \rangle = \overline{t}\,\phi(u)(x) \end{aligned}$$

for all $x \in U$. The map ϕ is injective because if $\phi(u_1) = \phi(u_2)$ then $\langle x - u_1 \rangle = \langle x, u_2 \rangle$ for all $x \in U$, so $u_1 = u_2$ by Theorem 5.8. To show that ϕ is surjective, let $g \in U^*$. We must find $u \in U$ so that $\phi(u) = g$, that is so that $\langle x, u \rangle = g(x)$ for all $x \in U$. Choose an orthonormal basis $\mathcal{A} = \{u_1, \cdots, u_n\}$ for U. In order to obtain $g(x) = \langle x, u \rangle$ for all $x \in U$, it suffices to have $g(u_k) = \langle u_k, u \rangle$ for all indices k. We choose $u = \sum_{i=1}^{n} \overline{g(u_i)}\, u_i$ so that $\langle u, u_k \rangle = \overline{g(u_k)}$ and then we have $g(u_k) = \overline{\langle u, u_k \rangle} = \langle u_k, u \rangle$, as required.

7.12 Note: Let W be an inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and let $U \subseteq W$ be a subspace. Then the above map $\phi_W : W \to W^*$ given by $\phi_W(w)(x) = \langle x, w \rangle$ sends $U^\perp$ to U°. Indeed for $x \in W$ we have

$$\begin{aligned} u \in U^\perp &\iff \langle u, x \rangle = 0 \text{ for all } x \in U \iff \langle x, u \rangle = 0 \text{ for all } x \in U \\ &\iff \phi_W(u)(x) = 0 \text{ for all } x \in U \iff \phi_W(u) \in U^\circ. \end{aligned}$$

7.13 Definition: Let U and V be finite dimensional inner product spaces over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let $L : U \to V$ be a linear map. The **adjoint** of L is the map $L^* : V \to U$ given by

$$L^* = {\phi_U}^{-1} \circ L^T \circ \phi_V .$$

7.14 Note: For a map $M : V \to U$, where U and V are finite dimensional inner product spaces over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, we have

$$\begin{aligned} M = L^* &\iff M = {\phi_U}^{-1} \circ L^T \circ \phi_V \iff \phi_U \circ M = L^T \circ \phi_V \\ &\iff \phi_U\big(M(y)\big) = L^T\big(\phi_V(y)\big) \text{ for all } y \in V \\ &\iff \phi_U\big(M(y)\big) = \phi_V(y) \circ L \text{ for all } y \in V \\ &\iff \phi_U\big(M(y)\big)(x) = \phi_V(y)\big(L(x)\big) \text{ for all } x \in U, y \in V \\ &\iff \big\langle x, M(y)\big\rangle = \big\langle L(x), y\big\rangle \text{ for all } x \in U, y \in V. \end{aligned}$$

Thus the adjoint of L is the unique map $L^* : V \to U$ with the property that

$$\big\langle L(x), y\big\rangle = \big\langle x, L^*(y)\big\rangle \text{ for all } x \in U, y \in V.$$

When $\mathbb{F} = \mathbb{R}$, the adjoint L^* is clearly a linear map because it is the composite of linear maps. When $\mathbb{F} = \mathbb{C}$, the adjoint L^* is again linear since the map L^T is linear and the maps ${\phi_U}^{-1}$ and ϕ_V are conjugate-linear. Indeed for $y \in V$ and $t \in \mathbb{C}$ we have

$${\phi_U}^{-1}\Big(L^T\big(\phi_V(ty)\big)\Big) = {\phi_U}^{-1}\Big(L^T\big(\overline{t}\,\phi_V(y)\big)\Big) = {\phi_U}^{-1}\Big(\overline{t}\,L^T\big(\phi_V(y)\big)\Big) = t\,{\phi_U}^{-1}\Big(L^T\big(\phi_V(y)\big)\Big).$$

7.15 Theorem: *Let U and V be finite dimensional inner product spaces over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and let $L : U \to V$ be a linear map. Let $\mathcal{A}$ and $\mathcal{B}$ be orthonormal bases for U and V. Then*

$$\big[L^*\big]_{\mathcal{A}}^{\mathcal{B}} = \Big([L]_{\mathcal{B}}^{\mathcal{A}}\Big)^* .$$

Proof: Let $\mathcal{A} = \{u_1, \cdots, u_k\}$ and $\mathcal{B} = \{v_1, \cdots, v_l\}$. Then

$$[L]_{\mathcal{B}}^{\mathcal{A}} = \Big(\big[L(u_1)\big]_{\mathcal{B}}, \cdots, \big[L(u_k)\big]_{\mathcal{B}}\Big) = \begin{pmatrix} \big\langle L(u_1), v_1\big\rangle & \big\langle L(u_2), v_1\big\rangle & \cdots & \big\langle L(u_k), v_1\big\rangle \\ \big\langle L(u_1), v_2\big\rangle & \big\langle L(u_2), v_2\big\rangle & \cdots & \big\langle L(u_k), v_2\big\rangle \\ \vdots & \vdots & & \vdots \\ \big\langle L(u_1), v_l\big\rangle & \big\langle L(u_2), v_l\big\rangle & \cdots & \big\langle L(u_k), v_l\big\rangle \end{pmatrix}$$

and

$$\big[L^*\big]_{\mathcal{A}}^{\mathcal{B}} = \Big(\big[L^*(v_1)\big]_{\mathcal{A}}, \cdots, \big[L^*(v_l)\big]_{\mathcal{A}}\Big) = \begin{pmatrix} \big\langle L^*(v_1), u_1\big\rangle & \big\langle L^*(v_2), u_1\big\rangle & \cdots & \big\langle L^*(v_l), u_1\big\rangle \\ \big\langle L^*(v_1), u_2\big\rangle & \big\langle L^*(v_2), u_2\big\rangle & \cdots & \big\langle L^*(v_l), u_2\big\rangle \\ \vdots & \vdots & & \vdots \\ \big\langle L^*(v_1), u_k\big\rangle & \big\langle L^*(v_2), u_k\big\rangle & \cdots & \big\langle L^*(v_l), u_k\big\rangle \end{pmatrix}$$

so the (i, j) entry of the matrix $\big[L^*\big]_{\mathcal{A}}^{\mathcal{B}}$ is

$$\Big(\big[L^*\big]_{\mathcal{A}}^{\mathcal{B}}\Big)_{i,j} = \big\langle L^*(v_j), u_i\big\rangle = \big\langle v_j, L(u_i)\big\rangle = \overline{\big\langle L(u_i), v_j\big\rangle} = \overline{\Big([L]_{\mathcal{B}}^{\mathcal{A}}\Big)}_{j,i}.$$

7.16 Remark: We now wish to extend our definition of the adjoint of a linear map $L : U \to V$ to include the case in which U and V are infinite dimensional.

7.17 Theorem: *Let U and V be inner product spaces over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and let $L : U \to V$ be a linear map. Suppose that there exists a map $M : V \to U$ with the property that*

$$\big\langle L(x), y\big\rangle = \big\langle x, M(y)\big\rangle \text{ for all } x \in U, y \in V.$$

Then M is unique and linear.

Proof: To prove that M is unique, suppose that another map $N : V \to U$ has the property that $\big\langle L(x), y\big\rangle = \big\langle x, N(y)\big\rangle$ for all $x \in U$ and $y \in V$. Then for all $y \in V$ we have $\langle x, M(y)\rangle = \langle L(x), y\rangle = \langle x, N(y)\rangle$ for all $x \in U$, and so $M(y) = N(y)$ by Theorem 5.8. Since $M(y) = N(y)$ for all $y \in V$, we have $M = N$. To see that M is linear, let $y, y_1, y_2 \in V$ and let $t \in \mathbb{F}$. Since

$$\begin{aligned}\big\langle x, M(y_1 + y_2)\big\rangle &= \langle L(x), y_1 + y_2\rangle = \langle L(x), y_1\rangle + \langle L(x), y_2\rangle \\ &= \langle x, M(y_1)\rangle + \langle x, M(y_2)\rangle = \big\langle x, M(y_1) + M(y_2)\big\rangle\end{aligned}$$

for all $x \in U$, we have $M(y_1 + y_2) = M(y_1) + M(y_2)$ by Theorem 5.8. Since

$$\langle x, M(ty)\rangle = \langle L(x), ty\rangle = \overline{t}\,\langle L(x), y\rangle = \overline{t}\,\langle x, M(y)\rangle = \langle x, tM(y)\rangle$$

for all $x \in U$, we have $M(ty) = tM(y)$ by Theorem 5.8. Thus M is linear.

7.18 Definition: Let U and V be inner product spaces over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and let $L : U \to V$ be a linear map. If there exists a map $L^* : V \to U$ with the property that

$$\big\langle L(x), y\big\rangle = \big\langle x, L^*(y)\big\rangle \text{ for all } x \in U, y \in V.$$

then, by the above theorem, L^* is unique and linear, and we call it the **adjoint** of L.