LECTURE 18

Adjoint maps

We finish our discussion of linear transformations by considering the case where the domain and codomain are inner product spaces. This context provides a rich source of additional properties that we can identify for our maps, with a wide range of applications.

In the following, we will assume that our vector spaces are real and finite-dimensional and are equipped with

- inner-products, generically denoted by $\langle \cdot, \cdot \rangle_V$;
- bases which are orthonormal with respect to $\langle \cdot, \cdot \rangle_V$.

We will also denote the standard Euclidean inner product on $\mathbb{R}^n$ by $\langle \cdot, \cdot \rangle_n$, or occasionally by $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$.

18.1 Adjoint maps

Let V, W be inner-product spaces and let $f \in \mathcal{L}(V, W)$. Let us consider defining another linear map $g : W \to V$ related to f in some sense. To narrow down this task, consider the following two results. The first is essentially Proposition 6.4, using notation relevant to the current topic.

Proposition 18.1 Let $(V, \langle \cdot, \cdot \rangle_V)$ be an inner product space. Let $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ denote a basis for V which is orthonormal with respect to $\langle \cdot, \cdot \rangle_V$. Then, for any $\mathbf{u}, \mathbf{v} \in V$,

$$\langle \mathbf{u}, \mathbf{v} \rangle_V = \langle \mathbf{x}, \mathbf{y} \rangle_n ,$$

where $\mathbf{u} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n$ and $\mathbf{v} = y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + \cdots + y_n\mathbf{v}_n$.

Proof. We have, using the linearity of the inner product and the orthonormality of $\mathbf{v}_i$,

$$\langle \mathbf{u}, \mathbf{v} \rangle_V = \langle x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n, y_1\mathbf{v}_1 + \cdots + y_n\mathbf{v}_n \rangle = x_1y_1 + \cdots + x_ny_n = \langle \mathbf{x}, \mathbf{y} \rangle_n .$$

■

The inner product on $\mathbb{R}^n$ can also be given the following convenient form involving the product of two 'matrices': the $1 \times n$ matrix (row vector) $\mathbf{x}^T \in \mathbb{R}^{1\times n}$ and the $n \times 1$ matrix (column vector) $\mathbf{y} \in \mathbb{R}^{n\times 1}$:

$$\langle \mathbf{x}, \mathbf{y} \rangle_n = \mathbf{x}^T\mathbf{y} = \mathbf{y}^T\mathbf{x}.$$

We used the *transpose* notation to denote the row vector, which is seen here as the transpose of the column vector $\mathbf{x}$. Generally, one can define the transpose A^T of a matrix as the matrix with entries $[A^T]_{ij} = [A]_{ji}$.

The following result is a corollary.

Proposition 18.2 Let $f \in \mathcal{L}(V, W)$, where V, W are inner product spaces equipped with orthonormal bases. Let $A \in \mathbb{R}^{m\times n}$ be the matrix representation of f with respect to these bases. Then

$$\langle f(\mathbf{v}), \mathbf{w}\rangle_W = \langle A\mathbf{x}, \mathbf{y}\rangle_m = \mathbf{y}^T A\mathbf{x},$$

where $\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n$ and $\mathbf{w} = y_1\mathbf{w}_1 + y_2\mathbf{w}_2 + \cdots + y_m\mathbf{w}_m$.

Proof. The result follows by applying Proposition 18.1 to the evaluation of the inner product $\langle \mathbf{u}, \mathbf{w}\rangle$ where $\mathbf{u} = f(\mathbf{v})$, with $\varphi_V(\mathbf{u}) = A\mathbf{x}$. ■

The expression for the inner product in Proposition 18.2 can be written also in the form

$$\mathbf{y}^T A\mathbf{x} = \mathbf{y}^T(A\mathbf{x}) = (A\mathbf{x})^T\mathbf{y} = \mathbf{x}^T A^T\mathbf{y}.$$

This suggests that we could consider another linear map with matrix representation $A^T \in \mathbb{R}^{n\times m}$. This is described in the following definition.

Definition 18.1 — Adjoint. Let $f \in \mathcal{L}(V, W)$, where V, W are inner product spaces. The adjoint of f is the map $f^* : W \to V$ defined via

$$\langle f^*(\mathbf{w}), \mathbf{v}\rangle_V = \langle \mathbf{w}, f(\mathbf{v})\rangle_W .$$

R Note that one could also provide the definition of f^* based on the above discussion, namely,

$$\langle f^*(\mathbf{w}), \mathbf{v}\rangle_V = \mathbf{x}^T A^T\mathbf{y} = \mathbf{y}^T A\mathbf{x} = \langle f(\mathbf{v}), \mathbf{w}\rangle_W .$$

However, the expression given in the definition is preferred, as it provides a reminder that f and f^* take arguments $\mathbf{v}$ and $\mathbf{w}$ from different vector spaces.

The concept of adjoint map is well-defined, due to the following uniqueness result.

Proposition 18.3 The map f^* is the unique map satisfying the relation given in Definition 18.1.

Proof. Exercise. ■

Exercise 18.1 Show that f^* as defined in Definition 18.1 is a linear map: $f^* \in \mathcal{L}(W, V)$.

By the above discussion, the following result holds.

Proposition 18.4 Let $f \in \mathcal{L}(V, W)$, where V, W are inner product spaces equipped with orthonormal bases. Let A be the matrix representation of f. Then the matrix representation of f^* is A^T.

We can actually derive an explicit expression for the action of the adjoint map on a vector $\mathbf{w} \in W$.

Proposition 18.5 Let $f \in \mathcal{L}(V, W)$, where V, W are inner product spaces equipped with orthonormal bases. Then the adjoint of f is given by

$$f^*(\mathbf{w}) = \sum_{i=1}^{n} \langle \mathbf{w}, f(\mathbf{v}_i)\rangle_W \, \mathbf{v}_i.$$

Proof. Recall first that orthonormal bases allow for a Fourier representation of a vector, with the coefficients written as inner products (see Lecture 8):

$$\mathbf{v} = \sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{v}_i \rangle_V \mathbf{v}_i.$$

Setting $\mathbf{v} = f^*(\mathbf{w})$, we find the coefficients take the form

$$\langle \mathbf{v}, \mathbf{v}_i \rangle_V = \langle f^*(\mathbf{w}), \mathbf{v}_i \rangle_V = \langle \mathbf{w}, f(\mathbf{v}_i) \rangle_W ,$$

where we used the definition of adjoint. ∎

Let us consider an example of derivation of an adjoint.

Example 18.1 Let $f : \mathbb{R}^4 \to \mathbb{R}^2$ be given via

$$f(\mathbf{x}) = f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) := \begin{bmatrix} x_1 + x_3 \\ x_2 + x_4 \end{bmatrix}.$$

We assume that $\mathbb{R}^4, \mathbb{R}^2$ are equipped with the standard Euclidean inner products, with respect to which the usual canonical bases are orthonormal. Using the expression given in the previous proposition, we find

$$f^*\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = \sum_{i=1}^{4} \langle \mathbf{y}, f(\mathbf{e}_i) \rangle_{\mathbb{R}^2} \mathbf{e}_i = \mathbf{y}^T \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mathbf{e}_1 + \mathbf{y}^T \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mathbf{e}_2 + \mathbf{y}^T \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mathbf{e}_3 + \mathbf{y}^T \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mathbf{e}_4 = \begin{bmatrix} y_1 \\ y_2 \\ y_1 \\ y_2 \end{bmatrix}.$$

Note that the same map is obtained if instead we derive the matrix representation of f and use its transpose to find the definition of f^*:

$$f(\mathbf{x}) = A\mathbf{x} \Longrightarrow f^*(\mathbf{y}) = A^T\mathbf{y},$$

where

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

Exercise 18.2 Find the adjoint map f^* from the previous example by using Definition 18.1) directly.

Proposition 18.6 The following properties hold for all $f, g : V \to W$, where V, W are real inner product spaces:

- $(f + g)^* = f^* + g^*$;
- $(af)^* = af^*$;
- $(f^*)^* = f$;
- $(f \circ g)^* = g^* \circ f^*$.

Proof. Exercise. ∎

We end the discussion on adjoint maps with the following result.

Proposition 18.7 Let $f : V \to W$, where V, W are inner-product spaces. Then $\ker f^* = (\mathrm{im}\ f)^\perp$.

Proof. Let $\mathbf{w} \in \ker f^*$. Then $f^*(\mathbf{w}) = \mathbf{0}_V$. Hence, for all $\mathbf{v} \in V$,

$$\langle f^*(\mathbf{w}), \mathbf{v} \rangle_V = 0 \iff \langle \mathbf{w}, f(\mathbf{v}) \rangle_W = 0 \iff \mathbf{w} \perp f(\mathbf{v}) \iff \mathbf{w} \in \text{im } f^\perp.$$

Hence, $\ker f^* = (\text{im } f)^\perp$. ∎

R The above result can be used to derive other similar identities, e.g., replacing f with f^*, we find $\ker f = (\text{im } f^*)^\perp$, while taking the orthogonal complement yields $\text{im } f = (\ker f^*)^\perp$ etc.

Corollary 18.8 Let $f : V \to W$, where V, W are inner product spaces. Then
i. $V = \text{im } f^* \oplus \ker f$;
ii. $W = \text{im } f \oplus \ker f^*$.

Proof. We use Proposition 18.7 (see also remark) and the concept of orthogonal decomposition of a vector space.
i. We find
$$V = \ker f \oplus (\ker f)^\perp = \ker f \oplus ((\text{im } f^*)^\perp)^\perp = \ker f \oplus \text{im } f^*$$
ii. Similarly,
$$W = \text{im } f \oplus (\text{im } f)^\perp = \text{im } f \oplus \ker f^*.$$
∎

Finally, we can combine these orthogonal decompositions with the rank-nullity formula to obtain the following result.

Proposition 18.9 Let $f : V \to W$, where V, W are inner product spaces. Then
$$\text{rank } f = \text{rank } f^*.$$

Proof. Since $V = \text{im } f^* \oplus \ker f$, we have
$$\dim V = \dim \text{im } f^* + \dim \ker f.$$
By the rank-nullity formula for $f : V \to W$, we have
$$\dim V = \dim \text{im } f + \dim \ker f.$$
Hence, $\dim \text{im } f = \dim \text{im } f^*$, i.e., $\text{rank } f = \text{rank } f^*$. ∎

An immediate corollary is the following well-known result.

Corollary 18.10 For any matrix there holds $\text{rank } A = \text{rank } A^T$.