Lecture 5
Adjoints of Linear Maps and The Spectral Theorem

5.1 The Dual Vector Space of a Vector Space

If V is a vector space then the vector space $L(V, \mathbf{R})$ of linear maps of V into the one-dimensional vector space of scalars, $\mathbf{R}$ plays an important role in many considerations. It is called the *dual space* of V and is denoted by V^*. Elements of V^* are called *linear functionals* on V.

5.1.1 Remark. If $v_1, \ldots, v_n$ is any basis for V, then (since 1 is a basis for $\mathbf{R}$) it follows from 4.1.1 that there is a uniquely determined basis $\ell_1, \ldots, \ell_n$ for V^* such that $\ell_i(v_j) = \delta^i_j$. This is called the *dual basis* to $v_1, \ldots, v_n$. In particular, V^* has the same dimension as V.

5.2 The Self-Duality of Inner-Product Spaces

If V is an inner-product space, then there is an important way to represent elements of V^*. Namely, recall that the inner-product is linear in each variable when the other variable is held fixed. This means that for each vector $v \in V$ we can define an element $v^* : V \to \mathbf{R}$ of V^* by the formula $v^*(x) := \langle x, v \rangle$. Since the inner product is linear in **both** variables, it follows that the map $v \mapsto v^*$ is linear. If $v^* = 0$, then in particular $0 = v^*(v) = \langle v, v \rangle$, so by positive definiteness $v = 0$, i.e., the kernel of the map $v \mapsto v^*$ is zero, and since V and V^* have the same dimension, it follows that this map is an isomorphism—i.e., every linear functional on an inner-product space V can be expressed uniquely as the inner-product with some fixed element of V. This is often expressed by saying that inner-product spaces are "self-dual".

5.2.1 Remark. Note that if e_i is an orthonormal basis for V, then e_i^* is the dual basis for V^*.

5.3 Adjoint Linear Maps

Now let V and W be finite dimensional inner-product spaces and $T : V \to W$ a linear map. We will next define a linear map $T^* : W \to V$ called the adjoint of T that satisfies the identity $\langle Tv, w \rangle = \langle v, T^*w \rangle$ for all $v \in V$ and all $w \in W$.

If we fix w in W, then to define T^*w we note that the map $v \to \langle Tv, w \rangle$ is clearly a linear functional on V, i.e., an element of V^*, so by self-duality there is a uniquely defined element T^*w in V such that $\langle v, T^*w \rangle = \langle Tv, w \rangle$ for all v in V.

We recapitulate the above as a formal definition.

5.3.1 Definition. Let V, W be finite dimensional inner-product spaces and $T \in L(V, W)$. The *adjoint* of T is the unique element $T^* \in L(W, V)$ satisfying the identity:

$$\langle v, T^* w \rangle = \langle Tv, w \rangle \, .$$

Here are some exercises involving adjoints and their basic properties.

▷ **5.3—Exercise 1.** Show that $(T^*)^* = T$.

▷ **5.3—Exercise 2.** Recall that if T_{ij} is an $m \times n$ matrix (i.e., m rows and n columns) and S_{ji} an $n \times m$ matrix, then S_{ji} is called the *transpose* of T_{ij} if $T_{ij} = S_{ji}$ for $1 \le i \le m$, $1 \le j \le n$. Show that if we choose orthonormal bases for V and W, then the matrix of T^* relative to these bases is the transpose of the matrix of T relative to the same bases.

▷ **5.3—Exercise 3.** Show that $\ker(T)$ and $\operatorname{im}(T^*)$ are orthogonal complements in V, and similarly, $\operatorname{im}(T)$ and $\ker(T^*)$ are each other's orthogonal complements in W. (Note that by Exercise 1, you only have to prove one of these.)

▷ **5.3—Exercise 4.** Show that a linear operator T on V is in the orthogonal group $\boldsymbol{O}(V)$ if and only if $TT^* = I$ (where I denotes the identity map of V) or equivalently, if and only if $T^* = T^{-1}$.

If $T : V \to V$ is a linear operator on V, then T^* is also a linear operator on V, so it makes sense to compare them and in particular ask if they are equal.

5.3.2 Definition. A linear operator on an inner-product space V is called *self-adjoint* if $T^* = T$, i.e., if $\langle Tv_1, v_2 \rangle = \langle v_1, Tv_2 \rangle$ for all $v_1, v_2 \in V$.

Note that by Exercise 3 above, self-adjoint operators are characterized by the fact that their matrices with respect to an orthonormal basis are symmetric.

▷ **5.3—Exercise 5.** Show that if W is a linear subspace of the inner-product space V, then the orthogonal projection P of V on W is a self-adjoint operator on V.

5.3.3 Definition. If T is a linear operator on V, then a linear subspace $U \subseteq V$ is called a T-invariant subspace if $T(U) \subseteq U$, i.e., if $u \in U$ implies $Tu \in U$.

5.3.4 Remark. Note that if U is a T-invariant subspace of V, then T can be regarded as a linear operator on U by restriction, and clearly if T is self-adjoint, so is its restriction.

▷ **5.3—Exercise 6.** Show that if $T : V \to V$ is a self-adjoint operator, and $U \subseteq V$ is a T-invariant subspace of V, the $U^\perp$ is also a T-invariant subspace of V.

5.4 Eigenvalues and Eigenvectors of a Linear Operator.

In this section, $T : V \to V$ is a linear operator on a real vector space V. If λ is a real number, then we define the linear subspace $E_\lambda(T)$ of V to be the set of $v \in V$ such that $Tv = \lambda v$. In other words, if I denotes the identity map of V, then $E_\lambda(T) = \ker(T - \lambda I)$.

Of course the zero vector is always in $E_\lambda(T)$. If $E_\lambda(T)$ contains a non-zero vector, then we say that λ is an *eigenvalue* of T and that $E_\lambda(T)$ is the λ-eigenspace of T. A non-zero vector in $E_\lambda(T)$ is called an *eigenvector* of T belonging to the eigenvector λ. The set of all eigenvalues of T is called the *spectrum* of T (the name comes from quantum mechanics) and it is denoted by $\mathrm{Spec}(T)$.

▷ **5.4—Exercise 1.** Show that a linear operator T on V has a diagonal matrix in a particular basis for V if and only if each element of the basis is an eienvector of T, and that then $\mathrm{Spec}(T)$ consists of the diagonal elements of the matrix.

The following is an easy but very important fact.

Theorem. *If T is a self-adjoint linear operator on an inner-product space V and λ_1, λ_2 are distinct real numbers, then $E_{\lambda_1}(T)$ and $E_{\lambda_2}(T)$ are orthogonal subspaces of V. In particular, eigenvectors of T that belong to different eigenvalues are orthogonal.*

▷ **5.4—Exercise 2.** Prove this theorem. (Hint: Let $v_i \in E_{\lambda_i}(T), i = 1, 2$. You must show that $\langle v_1, v_2 \rangle = 0$. Start with the fact that $\langle Tv_1, v_2 \rangle = \langle v_1, Tv_2 \rangle$.)

A general operator T on a vector space V need not have any eigenvalues, that is, $\mathrm{Spec}(T)$ may be empty. For example a rotation in the plane (by other than π or 2π radians) clearly has no eigenvectors and hence no eigenvalues. On the other hand self-adjoint operators always have at least one eigenvector. That is:

Spectral Lemma. *If V is an inner-product space of positive, finite dimension and if $T : V \to V$ is a self-adjoint operator on V, then* $\mathrm{Spec}(T)$ *is non-empty, i.e., T has at least one eigenvalue and hence at least one eigenvector.*

We will prove this result later after some preparation. But next we show how it leads to an easy proof of the extremely important:

Spectral Theorem for Self-Adjoint Operators. *If V is an inner inner-product space of posiive, finite dimension and $T : V \to V$ is a self-adjoint operator on V, then V has an orthonormal basis consisting of eigenvectors of T. In other words, T has a diagonal matrix in some orthonormal basis for V.*

PROOF. We prove this by induction on the dimension n of V. If $n = 1$ the theorem is trivial, since any non-zero vector in V is clearly an eigenvector, Thus we can assume that the theorem is valid for all self-adjoint operators on inner-product spaces of dimension less than n. By the Spectral Lemma, we can find at least one eigenvector w for T. Let $e_1 = w / \|w\|$ and let W be the one-dimensional space spanned by w. The fact that w is an eigenvector implies that W is a T-invariant linear subspace of V, and by 5.3, so is $W^\perp$. Since $W^\perp$ has dimension $n - 1$, by the inductive hypothesis T restricted to $W^\perp$ has an orthonormal basis $e_2, \ldots . e_n$ of eigenvectors of T, and then $e_1, \ldots, e_n$ is an orthonormal basis of V consisting of eigenvectors of T.

5.5 Finding One Eigenvector of a Self-Adjoint Operator

In this section we will outline the strategy for finding an eigenvector v for a self-adjoint operator T on an inner product space V. Actually carrying out this strategy will involve some further preparation. In particular, we will need first to review the basic facts about differential calculus in vector spaces (i.e., "multivariable calculus").

We choose an orthonormal basis $e_1, \ldots, e_n$ for V, and let T_{ij} denote the matrix of T in this basis.

We define a real-valued function F on V by $F(x) = \langle Tx, x\rangle$. F is called the *quadratic form on V defined by T*. The name comes from the following fact.

▷ **5.5—Exercise 1.** Show that if $x = x_1e_1 + \cdots + x_ne_n$, then $F(x) = \sum_{i,j=1}^n T_{ij}x_ix_j$, and using the symmetry of T_{ij} deduce that $\frac{\partial F}{\partial x_i} = 2\sum_{k=1}^n T_{ik}x_k$.

We will denote by $S(V)$ the set $\{x \in V \mid \|x\| = 1\}$, i.e., the "unit sphere" of normalized vectors. Note that if $x = x_1e_1 + \cdots + x_ne_n$, then $x \in S(V)$ if and only if $x_1^2 + \cdots + x_n^2 = 1$, so if we define $G : V \to \mathbf{R}$ by $G(x) := x_1^2 + \cdots + x_n^2 - 1$, then $S(V) = \{x \in V \mid G(x) = 0\}$.

5.5.1 Remark. Now $T_{ij} = \frac{\partial F}{\partial x_i} = 2\sum_{k=1}^n T_{ik}x_k$ is just twice the i-th component of Tx in the basis e_j. On the other hand $\frac{\partial G}{\partial x_i} = 2x_i$, which is twice the i-th component of x in this basis. It follows that a point x of $S(V)$ is an eigenvector of T provided there is a constant λ (the eigenvalue) such that $\frac{\partial F}{\partial x_i}(x) = \lambda\frac{\partial G}{\partial x_i}(x)$. If you learned about constrained extrema and Lagrange multipliers from advanced calculus, this should look familiar. Let me quote the relevant theorem,

Theorem on Constrained Extrema. *Let F and G be two differentiable real-valued functions on $\mathbf{R}^n$, and define a "constraint surface" $S \subseteq \mathbf{R}^n$ by $S = \{x \in V \mid G(x) = 0\}$. Let $p \in S$ be a point where F assumes its maximum value on S, and assume that the partial derivatives of G do not all vanish at p. Then the partial derivatives of F and of G are proportional at p, i.e., there is a constant λ (the "Lagrange Multiplier") such that $\frac{\partial F}{\partial x_i}(x) = \lambda\frac{\partial G}{\partial x_i}(x)$ for $i = 1, \ldots, n$.*

(We will sketch the proof of the Theorem on Constrained Extrema in the next lecture.)

Thus, **to find an eigenvector of T, all we have to do is choose a point of $S(V)$ where the quadratic form $F(x) = \langle Tx, x\rangle$ assumes its maximum value on $S(V)$.**