MATH 725 NOTES
STRUCTURE THEORY FOR NORMAL OPERATORS

MARIO L. GUTIERREZ ABED

Linear Functionals

Definition. *Let V be a vector space over $\mathbb{F}$. A linear transformation $f \in \mathcal{L}(V, \mathbb{F})$ whose values lie in the base field $\mathbb{F}$ is called a* ***linear functional*** *on V. The vector space of all linear functionals on V is denoted by V^* and is called the* ***algebraic dual space*** *of V.*[1] ★

<u>*Remark*</u>: For any $f \in V^*$, the rank-nullity theorem is

$$\dim(V) = \dim(\ker(f)) + \dim(\operatorname{Im}(f)).$$

But since $\operatorname{Im}(f) \subseteq \mathbb{F}$, we have either $\operatorname{Im}(f) = 0$ (in which case f is the zero linear functional) or $\operatorname{Im}(f) = \mathbb{F}$ (in which case f is surjective). In other words, a nonzero linear functional is surjective. Moreover, if $f \neq 0$, then

$$\operatorname{codim}(\ker(f)) = \dim(V/\ker(f)) = 1,$$

and if $\dim(V) < \infty$ then

$$\dim(\ker(f)) = \dim(V) - \dim(V/\ker(f)) = \dim(V) - 1.$$

Thus, in dimensional terms, the kernel of a linear functional is a very "large" subspace of the domain V.

The following theorem will prove very useful:

Theorem 1. *We have the following important facts about linear functionals:*

a) For any nonzero vector $v \in V$, there exists a linear functional $f \in V^*$ for which $f(v) \neq 0$.

b) A vector $v \in V$ is zero if and only if $f(v) = 0$ for all $f \in V^*$.

c) Let $f \in V^*$. If $f(x) \neq 0$, then

$$V = \langle x \rangle \oplus \ker(f).$$

d) Two nonzero linear functionals $f, g \in V^*$ have the same kernel if and only if there is a nonzero scalar λ such that $f = \lambda g$.

[1] The adjective *algebraic* is needed here, since there is another type of dual space that is defined on general normed vector spaces, where continuity of linear transformations makes sense. We will discuss these so-called *continuous dual spaces* later on.

Proof. We are proving parts *c)* and *d)*. For part *c)*, if $0 \neq v \in \langle x \rangle \cap \ker(f)$, then $f(v) = 0$ and $v = ax$ for $0 \neq a \in \mathbb{F}$, from which we have that $f(x) = 0$, which is false. Hence $\langle x \rangle \cap \ker(f) = \{0\}$ and the direct sum $S = \langle x \rangle \oplus \ker(f)$ exists. Also, for any $v \in V$ we have

$$v = \frac{f(v)}{f(x)}x + \left(v - \frac{f(v)}{f(x)}x\right) \in \langle x \rangle + \ker(f)$$

and so $V = \langle x \rangle \oplus \ker(f)$.

For part *d)*, if $f = \lambda g$ for $\lambda \neq 0$, then $\ker(f) = \ker(g)$. Conversely, if $K = \ker(f) = \ker(g)$, then for $x \notin K$ we have by part c) that

$$V = \langle x \rangle \oplus K.$$

Of course, $f|_K = \lambda g|_K$ for any λ. Therefore, if $\lambda = f(x)/g(x)$, it follows that $\lambda g(x) = f(x)$ and hence $f = \lambda g$. □

Dual Basis

Let V be a vector space with basis $\mathcal{B} = \{v_i \mid i \in I\}$. For each $i \in I$, we can define a linear functional $v_i^* \in V^*$, by the orthogonality condition

$$v_i^*(v_j) = \delta_{i,j}.$$

This brings us to the following theorem:

Theorem 2. *Let V be a vector space with basis $\mathcal{B} = \{v_i \mid i \in I\}$.*

- *a) The set $\mathcal{B}^* = \{v_i^* \mid i \in I\}$ is linearly independent.*
- *b) If V is finite-dimensional then $\mathcal{B}^*$ is a basis for V^*, called the* ***dual basis*** *of $\mathcal{B}$.*

Proof of a). Notice that by applying the equation

$$0 = a_{i_1}v_{i_1}^* + \cdots + a_{i_n}v_{i_n}^*$$

to the basis vector $v_{i_k} \in \mathcal{B}$, we get

$$0 = \sum_{j=1}^{k} a_{i_j}v_{i_j}^*(v_{i_k}) = \sum_{j=1}^{k} a_{i_j}\delta_{i_j,i_k} = a_{i_k} \qquad \text{for all } i_k.$$

□

Proof of b). Note that for any $f \in V^*$ we have

$$\sum_j f(v_j)v_j^*(v_i) = \sum_j f(v_j)\delta_{i,j} = f(v_i),$$

and so $f = \sum_j f(v_j)v_j^*$ is in the span of $\mathcal{B}^*$. By part *a)* we already know that $\mathcal{B}^*$ is linearly independent. Hence, $\mathcal{B}^*$ is a basis for V^*. □

Corollary 1. *If $\dim V < \infty$, then $\dim V^* = \dim V$.*

Remark: The functions $f \in V^*$ are defined on vectors in V, but we may also define f on subsets M of V by letting

$$f(M) = \{f(v) \mid v \in M\}.$$

Definition. *Let M be a nonempty subset of a vector space V. The* ***annihilator*** M^0 *of M is given by*

$$M^0 = \{f \in V^* \mid f(M) = \{0\}\}. \qquad \bigstar$$

The Riesz Representation Theorem

If x is a vector in an inner product space V, then the function $\phi_x \colon V \to \mathbb{F}$ defined by

$$\phi_x(v) = \langle v, x\rangle$$

is easily seen to be a linear functional on V. The following theorem shows that all linear functionals on a finite-dimensional inner product space V have this form. Then we will show that in the infinite-dimensional case, all continuous linear functionals on V have this form:

Theorem 3 (**Riesz Representation Theorem**). *Let V be a finite-dimensional inner product space and let $f \in V^*$ be a linear functional on V. Then there exists a unique vector $x \in V$ for which*

$$f(v) = \langle v, x\rangle \qquad \forall\, v \in V.$$

Now in the general case, we have the remarkable fact that every continuous linear functional on a Hilbert space arises as an inner product, as stated by the following theorem:

Theorem 4 (**Riesz Representation Theorem**). *Let ℓ be a continuous linear functional on a Hilbert space $\mathcal{H}$. Then, there exists a unique $g \in \mathcal{H}$ such that*

$$\ell(f) = \langle f, g\rangle \qquad \forall\, f \in \mathcal{H}.$$

Moreover, $\|\ell\| = \|g\|$.

Proof. Consider the subspace of $\mathcal{H}$ defined by

$$\mathcal{S} = \{f \in \mathcal{H} \mid \ell(f) = 0\}.$$

Since ℓ is continuous, the subspace $\mathcal{S}$, which is called the nullspace of ℓ, is closed. If $\mathcal{S} = \mathcal{H}$, then $\ell = 0$ and we take $g = 0$. Otherwise $\mathcal{S}^\perp$ is non-trivial and we may pick any $h \in \mathcal{S}^\perp$ with $\|h\| = 1$. With this choice of h we determine g by setting $g = \overline{\ell(h)}h$. Thus if we let $u = \ell(f)h - \ell(h)f$, then $u \in \mathcal{S}$, and therefore $\langle u, h\rangle = 0$. Hence

$$0 = \langle \ell(f)h - \ell(h)f, h\rangle = \ell(f)\langle h, h\rangle - \langle f, \overline{\ell(h)}h\rangle.$$

Since $\langle h, h\rangle = 1$, we find that $\ell(f) = \langle f, g\rangle$ as desired. $\square$

The first application of the Riesz representation theorem is to determine the existence of the "adjoint" of a linear transformation:

Theorem 5. *Let V and W be finite-dimensional inner product spaces over $\mathbb{F}$ and let and let $T \in \mathcal{L}(V, W)$. Then there is a unique function $T^*: W \to V$ that satisfies*

$$\langle Tv, w\rangle = \langle v, T^* w\rangle \qquad \forall\ v \in V, w \in W.$$

This function is called the ***adjoint***[2] *of T.*

Proof. For a fixed $w \in W$, consider the function $\theta_w : V \to \mathbb{F}$ defined by

$$\theta_w(v) = \langle T(v), w\rangle.$$

It is easy to verify that θ_w is a linear functional on V and so, by the *Riesz Representation Theorem*, there exists a unique vector $x \in V$ for which

$$\theta_w(v) = \langle v, x\rangle \qquad \text{for all } v \in V.$$

Hence, if $T^*(w) = x$ then

$$\langle T(v), w\rangle = \langle v, T^*(w)\rangle \qquad \text{for all } v \in V.$$

This establishes the existence and uniqueness of T^*.

To show that T^* is linear, observe that

$$\begin{aligned}
\langle v, T^*(\alpha w + \beta u)\rangle &= \langle T(v), \alpha w + \beta u\rangle \\
&= \bar{\alpha}\langle T(v), w\rangle + \bar{\beta}\langle T(v), u\rangle \\
&= \bar{\alpha}\langle v, T^*(w)\rangle + \bar{\beta}\langle v, T^*(u)\rangle \\
&= \langle v, \alpha T^*(w)\rangle + \langle v, \beta T^*(u)\rangle \\
&= \langle v, \alpha T^*(w) + \beta T^*(u)\rangle
\end{aligned}$$

for all $v \in V$, and so

$$T^*(\alpha w + \beta u) = \alpha T^*(w) + \beta T^*(u).$$

Hence $T^* \in \mathcal{L}(W, V)$. ☐

Example: Let's work out an example of how the adjoint is computed.

Define $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ by:

$$T(x_1, x_2, x_3) = (x_2 + 3x_3, 2x_1).$$

Thus T^* will be a function from $\mathbb{R}^2$ to $\mathbb{R}^3$. To compute T^*, fix a point $(y_1, y_2) \in \mathbb{R}^2$. Then

$$\begin{aligned}
\langle (x_1, x_2, x_3), T^*(y_1, y_2)\rangle &= \langle T(x_1, x_2, x_3), (y_1, y_2)\rangle \\
&= \langle (x_2 + 3x_3, 2x_1), (y_1, y_2)\rangle \\
&= x_2 y_1 + 3x_3 y_1 + 2x_1 y_2 \\
&= \langle (x_1, x_2, x_3), (2y_2, y_1, 3y_1)\rangle
\end{aligned}$$

for all $(x_1, x_2, x_3) \in \mathbb{R}^3$.

This shows that $T^*(y_1, y_2) = (2y_2, y_1, 3y_1)$.

[2]The word *adjoint* has another meaning in linear algebra, which is related to inverses. Be warned that the two meanings for adjoint are unrelated to one another.

Here are some of the basic properties of the adjoint:

Theorem 6. *Let V and W be finite-dimensional inner product spaces. For every $\sigma, \tau \in \mathcal{L}(V, W)$, and $\alpha \in F$, we have*

- $(\sigma + \tau)^* = \sigma^* + \tau^*$.
- $(\alpha\tau)^* = \bar{\alpha}\tau^*$.
- $\tau^{**} = \tau$.
- *If $V = W$, then $(\sigma\tau)^* = \tau^*\sigma^*$.*
- *If τ is invertible, then $(\tau^{-1})^* = (\tau^*)^{-1}$.*
- *If $V = W$ and $p(x) \in \mathbb{R}[x]$, then $p(\tau)^* = p(\tau^*)$.*

Now let us relate the kernel and image of a linear transformation to those of its adjoint:

Theorem 7. *Let V and W be finite-dimensional inner product spaces and let $\tau \in \mathcal{L}(V, W)$. Then we have*

- $\ker(\tau^*) = \operatorname{Im}(\tau)^{\perp}$.
- $\operatorname{Im}(\tau^*) = \ker(\tau)^{\perp}$.
- *τ is injective $\iff$ τ^* is surjective.*
- *τ is surjective $\iff$ τ^* is injective.*
- $\ker(\tau^*\tau) = \ker(\tau)$
- $\ker(\tau\tau^*) = \ker(\tau^*)$
- $\operatorname{Im}(\tau^*\tau) = \operatorname{Im}(\tau^*)$
- $\operatorname{Im}(\tau\tau^*) = \operatorname{Im}(\tau)$

Definition. *A linear operator τ on an inner product space is said to be **normal** if it commutes with its adjoint. That is, if*

$$\tau\tau^* = \tau^*\tau. \qquad \bigstar$$

Definition. *Let V be an inner product space and let $\tau \in \mathcal{L}(V)$. Then*

- *τ is **self-adjoint** (also called **Hermitian** in the complex case and **symmetric** in the real case), if*
$$\tau^* = \tau.$$
- *τ is called **skew-Hermitian** in the complex case and **skew-symmetric** in the real case, if*
$$\tau^* = -\tau.$$
- *τ is called **unitary** in the complex case and **orthogonal** in the real case if τ is invertible and*
$$\tau^* = \tau^{-1}. \qquad \bigstar$$

Theorem 8. *Let $\mathcal{H}$ be the set of self-adjoint operators on a finite-dimensional inner product space V. Then $\mathcal{H}$ satisfies the following properties:*

- **(Closure under addition)**
$$\sigma, \tau \in \mathcal{H} \implies \sigma + \tau \in \mathcal{H}.$$
- **(Closure under real scalar multiplication)**
$$\alpha \in \mathbb{R}, \tau \in \mathcal{H} \implies \alpha\tau \in \mathcal{H}.$$
- **(Closure under multiplication if the factors commute)**
$$\sigma, \tau \in \mathcal{H}, \sigma\tau = \tau\sigma \implies \sigma\tau \in \mathcal{H}.$$
- **(Closure under inverses)**
$$\tau \in \mathcal{H}, \tau \text{ is invertible} \implies \tau^{-1} \in \mathcal{H}.$$
- **(Closure under real polynomials)**
$$\tau \in \mathcal{H},\, p(\tau) \in \mathcal{H} \quad \text{for any} \quad p(x) \in \mathbb{R}[x].$$
- *A complex operator τ is Hermitian if and only if $\langle \tau(v), v\rangle$ (this is called the* ***quadratic form*** *of τ) is real for all $v \in V$.*
- *If $\mathbb{F} = \mathbb{C}$, or if $\mathbb{F} = \mathbb{R}$ and τ is symmetric, then $\tau = 0$ if and only if $\langle \tau(v), v\rangle = 0$.*
- *If τ is self-adjoint, then the characteristic polynomial of τ splits over $\mathbb{R}$ and so all complex eigenvalues are real.*

Theorem 9 (**The Structure Theorem for Normal Operators**)**.**

i) **(Complex Case)** *Let V be a finite-dimensional complex inner product space.*
   - *A linear operator τ on V is normal if and only if V has an orthonormal basis $\mathcal{B}$ consisting entirely of eigenvectors of τ; that is*
   $$V_\tau = \mathcal{E}_{\lambda_1} \odot \cdots \odot \mathcal{E}_{\lambda_k},$$
   where $\{\lambda_1, \dots, \lambda_k\}$ is the spectrum of τ. Put another way, τ is normal if and only if it is unitarily diagonalizable.
   - *Among the normal operators, the Hermitian operators are precisely those for which all complex eigenvalues are real.*
   - *Among the normal operators, the unitary operators are precisely those for which all eigenvalues have norm 1.*

ii) **(Real Case)** *Let V be a finite-dimensional real inner product space.*
   - *A linear operator τ on V is normal if and only if*
   $$V = \mathcal{E}_{\lambda_1} \odot \cdots \odot \mathcal{E}_{\lambda_k} \odot S_1 \odot \cdots \odot S_m,$$
   where $\{\lambda_1, \dots, \lambda_k\}$ is the spectrum of τ and each S_j is a two-dimensional τ-invariant subspace for which there exists an ordered basis $\mathcal{B}_j = (u_j, v_j)$ for which
   $$[\tau]_{\mathcal{B}_j} = \begin{bmatrix} a_j & -b_j \\ b_j & a_j \end{bmatrix}$$
   for $a_j, b_j \in \mathbb{R}$.

– *Among the real normal operators, the symmetric operators are precisely those for which there are no subspaces* U_i *in the decomposition of* V *above. Hence, an operator is symmetric if and only if it is orthogonally diagonalizable.*

– *Among the real normal operators, the orthogonal operators are precisely those for which the eigenvalues are equal to* ± 1 *and the matrices* $[\tau]_{\mathcal{B}_i}$ *described above have rows (and columns) of norm* 1*, that is,*

$$[\tau]_{\mathcal{B}_i} = \begin{bmatrix} \sin\theta & -\cos\theta \\ \cos\theta & \sin\theta \end{bmatrix}$$

for some $\theta \in \mathbb{R}$*.*