25. Duals, naturality, bilinear forms

1. *Dual vector spaces*

A **(linear) functional** $\lambda : V \longrightarrow k$ on a vector space V over k is a linear map from V to the field k itself, viewed as a one-dimensional vector space over k. The collection V^* of all such linear functionals is the **dual space** of V.

[1.0.1] **Proposition:** The collection V^* of linear functionals on a vector space V over k is itself a vector space over k, with the addition

$$(\lambda + \mu)(v) = \lambda(v) + \mu(v)$$

and scalar multiplication

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

Proof: The 0-vector in V^* is the linear functional which sends every vector to 0. The additive inverse $-\lambda$ is defined by

$$(-\lambda)(v) = -\lambda(v)$$

The distributivity properties are readily verified:

$$(\alpha(\lambda + \mu))(v) = \alpha(\lambda + \mu)(v) = \alpha(\lambda(v) + \mu(v)) = \alpha\lambda(v) + \alpha\mu(v) = (\alpha\lambda)(v) + (\alpha\mu)(v)$$

and

$$((\alpha + \beta) \cdot \lambda)(v) = (\alpha + \beta)\lambda(v) = \alpha\lambda(v) + \beta\lambda(v) = (\alpha\lambda)(v) + (\beta\lambda)(v)$$

as desired. ///

Let V be a finite-dimensional [1] vector space, with a basis $e_1, \ldots, e_n$ for V. A **dual basis** $\lambda_1, \ldots, \lambda_n$ for V^* (and $\{e_i\}$) is a basis for V^* with the property that

$$\lambda_j(e_i) = \begin{cases} 1 & (\text{for } i = j) \\ 0 & (\text{for } i \neq j) \end{cases}$$

From the definition alone it is not at all clear that a dual basis exists, but the following proposition proves that it does.

[1.0.2] **Proposition:** The dual space V^* to an n-dimensional vector space V (with n a positive integer) is also n-dimensional. Given a basis $e_1, \ldots, e_n$ for V, there exists a unique corresponding **dual basis** $\lambda_1, \ldots, \lambda_n$ for V^*, namely a basis for V^* with the property that

$$\lambda_j(e_i) = \begin{cases} 1 & (\text{for } i = j) \\ 0 & (\text{for } i \neq j) \end{cases}$$

Proof: Proving the existence of a dual basis corresponding to the given basis will certainly prove the dimension assertion. Using the *uniqueness* of expression of a vector in V as a linear combination of the basis vectors, we can unambiguously define a linear functional λ_j by

$$\lambda_j \left(\sum_i c_i e_i \right) = c_j$$

These functionals certainly have the desired relation to the basis vectors e_i. We must prove that the λ_j are a basis for V^*. If

$$\sum_j b_j \lambda_j = 0$$

then apply this functional to e_i to obtain

$$b_i = \left(\sum_j b_j \lambda_j \right) (e_i) = 0(e_i) = 0$$

This holds for every index i, so all coefficients are 0, proving the linear independence of the λ_j. To prove the spanning property, let λ be an arbitrary linear functional on V. We claim that

$$\lambda = \sum_j \lambda(e_j) \cdot \lambda_j$$

Indeed, evaluating the left-hand side on $\sum_i a_i e_i$ gives $\sum_i a_i \lambda(e_i)$, and evaluating the right-hand side on $\sum_i a_i e_i$ gives

$$\sum_j \sum_i a_i \, \lambda(e_j) \, \lambda_j(e_i) = \sum_i a_i \lambda(e_i)$$

since $\lambda_j(e_i) = 0$ for $i \neq j$. This proves that any linear functional is a linear combination of the λ_j. ///

Let W be a subspace of a vector space V over k. The **orthogonal complement** $W^\perp$ of W in V^* is

$$W^\perp = \{\lambda \in V^* : \lambda(w) = 0, \text{ for all } w \in W\}$$

[1] Some of the definitions and discussion here make sense for infinite-dimensional vector spaces V, but many of the conclusions are either false or require substantial modification to be correct. For example, by contrast to the proposition here, for infinite-dimensional V the (infinite) dimension of V^* is strictly larger than the (infinite) dimension of V. Thus, for example, the natural inclusion of V into its second dual V^{**} would fail to be an isomorphism.

• The orthogonal complement $W^\perp$ of a subspace W of a vector space V is a vector subspace of V^*.

Proof: Certainly $W^\perp$ contains 0. If $\lambda(w) = 0$ and $\mu(w) = 0$ for all $w \in W$, then certainly $(\lambda + \mu)(w) = 0$. Likewise, $(-\lambda)(w) = \lambda(-w)$, so $W^\perp$ is a subspace. ///

[1.0.3] **Corollary:** Let W be a subspace of a finite-dimensional vector space V over k.

$$\dim W + \dim W^\perp = \dim V$$

Proof: Let $e_1, \ldots, e_m$ be a basis of W, and extend it to a basis $e_1, \ldots, e_m$, $f_{m+1}, \ldots, f_n$ of V. Let $\lambda_1, \ldots, \lambda_m, \mu_{m+1}, \ldots, \mu_n$ be the corresponding dual basis of V^*. To prove the corollary it would suffice to prove that $\mu_{m+1}, \ldots, \mu_n$ form a basis for $W^\perp$. First, these functionals do lie in $W^\perp$, since they are all 0 on the basis vectors for W. To see that they span $W^\perp$, let

$$\lambda = \sum_{1 \le i \le m} a_i \lambda_i + \sum_{m+1 \le j \le n} b_j \mu_j$$

be a functional in $W^\perp$. Evaluating both sides on $e_\ell \in W$ gives

$$0 = \lambda(e_\ell) = \sum_{1 \le i \le m} a_i \lambda_i(e_\ell) + \sum_{m+1 \le j \le n} b_j \mu_j(e_\ell) = a_\ell$$

by the defining property of the dual basis. That is, every functional in $W^\perp$ is a linear combination of the μ_j, and thus the latter form a basis for $W^\perp$. Then

$$\dim W + \dim W^\perp = m + (n - m) = n = \dim V$$

as claimed. ///

The **second dual** V^{**} of a vector space V is the dual of its dual. There is a natural vector space homomorphism $\varphi : V \longrightarrow V^{**}$ of a vector space V to its second V^{**} by [2]

$$\varphi(v)(\lambda) = \lambda(v) \qquad (\text{for } v \in V,\ \lambda \in V^*)$$

[1.0.4] **Corollary:** Let V be a finite-dimensional vector space. Then the natural map of V to V^{**} is an isomorphism.

Proof: If v is in the kernel of the linear map $v \longrightarrow \varphi(v)$, then $\varphi(v)(\lambda) = 0$ for all λ, so $\lambda(v) = 0$ for all λ. But if v is non-zero then v can be part of a basis for V, which has a dual basis, among which is a functional λ such that $\lambda(v) = 1$. Thus, for $\varphi(v)(\lambda)$ to be 0 for all λ it must be that $v = 0$. Thus, the kernel of φ is $\{0\}$, so (from above) φ is an injection. From the formula

$$\dim \ker \varphi + \dim \operatorname{Im} \varphi = \dim V$$

[2] The austerity or starkness of this map is very different from formulas written in terms of matrices and column or row vectors. Indeed, this is a different sort of assertion. Further, the sense of *naturality* here might informally be construed exactly as that the formula does *not* use a basis, matrices, or any other manifestation of choices. Unsurprisingly, but unfortunately, very elementary mathematics does not systematically present us with good examples of naturality, since the emphasis is more often on computation. Indeed, we often take for granted the idea that two different sorts of computations will ineluctably yield the same result. Luckily, this is often the case, but becomes increasingly less obvious in more complicated situations.

it follows that $\dim \mathrm{Im}\,\varphi = \dim V$. We showed above that the dimension of V^* is the same as that of V, since V is finite-dimensional. Likewise, the dimension of $V^{**} = (V^*)^*$ is the same as that of V^*, hence the same as that of V. Since the dimension of the image of φ in V^{**} is equal to the dimension of V, which is the same as the dimension of V^{**}, the image must be all of V^{**}. Thus, $\varphi : V \longrightarrow V^{**}$ is an isomorphism. ///

[1.0.5] **Corollary:** Let W be a subspace of a finite-dimensional vector space V over k. Let $\varphi : V \longrightarrow V^{**}$ be the isomorphism of the previous corollary. Then

$$(W^\perp)^\perp = \varphi(W)$$

Proof: First, show that

$$\varphi(W) \subset (W^\perp)^\perp$$

Indeed, for $\lambda \in W^\perp$,

$$\varphi(w)(\lambda) = \lambda(w) = 0$$

On the other hand,

$$\dim W + \dim W^\perp = \dim V$$

and likewise

$$\dim W^\perp + \dim (W^\perp)^\perp = \dim V^* = \dim V$$

Thus, $\varphi(W) \subset (W^\perp)^\perp$ and

$$\dim (W^\perp)^\perp = \dim \varphi(W)$$

since φ is an isomorphism. Therefore, $\varphi(W) = (W^\perp)^\perp$. ///

As an illustration of the efficacy of the present viewpoint, we can prove a useful result about matrices.

[1.0.6] **Corollary:** Let M be an m-by-n matrix with entries in a field k. Let R be the subspace of k^n spanned by the rows of M. Let C be the subspace of k^m spanned by the columns of M. Let

$$\begin{array}{ccc} \text{column rank of } M & = & \dim C \\ \text{row rank of } M & = & \dim R \end{array}$$

Then

$$\text{column rank of } M = \text{row rank of } M$$

Proof: The matrix M gives a linear transformation $T : k^n \longrightarrow k^m$ by $T(v) = Mv$ where v is a column vector of length n. It is easy to see that the column space of M is the image of T. It is a little subtler that the row space is $(\ker T)^\perp$. From above,

$$\dim \ker T + \dim \mathrm{Im}\, T = \dim V$$

and also

$$\dim \ker T + \dim (\ker T)^\perp = \dim V$$

Thus,

$$\text{column rank } M = \dim \mathrm{Im}\, T = \dim (\ker T)^\perp = \text{ row rank } M$$

as claimed. ///

2. *First example of naturality*

We have in hand the material to illustrate a simple case of a **natural isomorphism** versus not-natural isomorphisms. This example could be given in the context of *category theory*, and in fact could be a first example, but it is possible to describe the phenomenon without the larger context. [3]

Fix a field k, and consider the map [4]

$$D : \{k\text{-vectorspaces}\} \longrightarrow \{k\text{-vectorspaces}\}$$

from the class of k-vectorspaces to itself given by **duality**, namely [5]

$$DV = V^* = \mathrm{Hom}_k(V, k)$$

Further, for a k-vectorspace homomorphism $f : V \longrightarrow W$ we have an associated map [6] f^* of the duals spaces

$$f^* : W^* \longrightarrow V^* \quad \text{by} \quad f^*(\mu)(v) = \mu(fv) \quad \text{for} \quad \mu \in W^*,\ v \in V$$

Note that f^* **reverses** direction, going from W^* to V^*, while the original f goes from V to W.

The map [7] $F = D \circ D$ associating to vector spaces V their *double* duals V^{**} also gives maps

$$f^{**} : V^{**} \longrightarrow W^{**}$$

for any k-vectorspace map $f : V \longrightarrow W$. (The direction of the arrows has been reversed twice, so is back to the original direction.)

And for each k-vectorspace V we have a k-vectorspace map [8]

$$\eta_V : V \longrightarrow V^{**} = (V^*)^*$$

given by

$$\eta_V(v)(\lambda) = \lambda(v)$$

The aggregate η of all the maps $\eta_V : V \longrightarrow V^{**}$ is a **natural transformation** [9] meaning that for all k-vectorspace maps

$$f : V \longrightarrow W$$

[3] Indeed, probably a collection of such examples should precede a development of general category theory, else there is certainly insufficient motivation to take the care necessary to develop things in great generality.

[4] In category-theory language a map on objects *and* on the maps among them is a **functor**. We will not emphasize this language just now.

[5] Certainly this class is not a *set*, since it is far too large. This potentially worrying foundational point is another feature of nascent category theory, as opposed to development of mathematics based as purely as possible on set theory.

[6] We might write $Df : DW \longrightarrow DV$ in other circumstances, in order to emphasize the fact that D maps both objects and the homomorphisms among them, but at present this is not the main point.

[7] Functor.

[8] The austere or stark nature of this map certainly should be viewed as being in extreme contrast to the coordinate-based linear maps encountered in introductory linear algebra. The very austerity itself, while being superficially simple, may cause some vertigo or cognitive dissonance for those completely unacquainted with the possibility of writing such things. Rest assured that this discomfort will pass.

[9] We should really speak of a natural transformation η *from* a *functor to another functor*. Here, η is from the identity functor on k-vectorspaces (which associates each V to itself), *to* the functor that associates to V its second dual V^{**}.

the diagram

$$\begin{array}{ccc} V & \xrightarrow{\eta_V} & V^{**} \\ f\downarrow & & \downarrow f^{**} \\ W & \xrightarrow{\eta_W} & W^{**} \end{array}$$

commutes. The commutativity of the diagram involving a particular M and N is called **functoriality** in M and in N. That the diagram commutes is verified very simply, as follows. Let $v \in V$, $\mu \in W^*$. Then

$$\begin{array}{rcll} ((f^{**} \circ \eta_V)(v))\,(\mu) & = & (f^{**}(\eta_V v))\,(\mu) & \text{(definition of composition)} \\ & = & (\eta_V v)(f^* \mu) & \text{(definition of } f^{**}) \\ & = & (f^*\mu)(v) & \text{(definition of } \eta_V) \\ & = & \mu(fv) & \text{(definition of } f^*) \\ & = & (\eta_W(fv))(\mu) & \text{(definition of } \eta_W) \\ & = & ((\eta_W \circ f)(v))\,(\mu) & \text{(definition of composition)} \end{array}$$

Since equality of elements of W^{**} is implied by equality of values on elements of W^*, this proves that the diagram commutes.

Further, for V finite-dimensional, we have

$$\dim_k V = \dim_k V^* = \dim_k V^{**}$$

which implies that each η_V must be an *isomorphism*. Thus, the aggregate η of the isomorphisms $\eta_V : V \longrightarrow V^{**}$ is called a **natural equivalence**. [10]

---

3. *Bilinear forms*

Abstracting the notion of **inner product** or **scalar product** or **dot product** on a vector space V over k is that of **bilinear form** or **bilinear pairing**. For purpose of *this* section, a bilinear form on V is a k-valued function of two V-variables, written $v \cdot w$ or $\langle v, w\rangle$, with the following properties for $u, v, w \in V$ and $\alpha \in k$
• (Linearity in both arguments) $\langle \alpha u + v, w\rangle = \alpha\langle u, w\rangle + \langle v, w\rangle$ and $\langle \alpha u, \beta v + v'\rangle = \beta\langle u, v\rangle + \langle u, v'\rangle$
• (Non-degeneracy) For all $v \neq 0$ in V there is $w \in V$ such that $\langle v, w\rangle \neq 0$. Likewise, for all $w \neq 0$ in V there is $v \in V$ such that $\langle v, w\rangle \neq 0$.
The two linearity conditions together are **bilinearity**.

In some situations, we may also have
• (Symmetry) $\langle u, v\rangle = \langle v, u\rangle$ However, the symmetry condition is not necessarily critical in many applications.

[3.0.1] **Remark:** When the scalars are the complex numbers $\mathbb{C}$, sometimes a variant of the symmetry condition is useful, namely a *hermitian* condition that $\langle u, v\rangle = \overline{\langle v, u\rangle}$ where the bar denotes complex conjugation.

[3.0.2] **Remark:** When the scalars are real or complex, sometimes, but not always, the non-degeneracy and symmetry are usefully replaced by a *positive-definiteness* condition, namely that $\langle v, v\rangle \geq 0$ and is 0 only for $v = 0$.

When a vector space V has a non-degenerate bilinear form $\langle,\rangle$, there are two natural linear maps $v \longrightarrow \lambda_v$ and $v \longrightarrow \mu_v$ from V to its dual V^*, given by

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

---

[10] More precisely, on the category of finite-dimensional k-vectorspaces, η is a natural equivalence of the identity functor with the second-dual functor.

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

That λ_v and μ_v are linear functionals on V is an immediate consequence of the linearity of $\langle,\rangle$ in its arguments, and the linearity of the map $v \longrightarrow \lambda_v$ itself is an immediate consequence of the linearity of $\langle,\rangle$ in its arguments.

[3.0.3] **Remark:** All the following assertions for $L : v \longrightarrow \lambda_v$ have completely analogous assertions for $v \longrightarrow \mu_v$, and we leave them to the reader.

[3.0.4] **Corollary:** Let V be a finite-dimensional vector space with a non-degenerate bilinear form $\langle,\rangle$. The linear map $L : v \longrightarrow \lambda_v$ above is an isomorphism $V \longrightarrow V^*$.

Proof: The non-degeneracy means that for $v \neq 0$ the linear functional λ_v is not 0, since there is $w \in V$ such that $\lambda_v(w) \neq 0$. Thus, the linear map $v \longrightarrow \lambda_v$ has kernel $\{0\}$, so $v \longrightarrow \lambda_v$ is *injective*. Since V is finite-dimensional, from above we know that it and its dual have the same dimension. Let $L(v) = \lambda_v$. Since

$$\dim \operatorname{Im} L + \dim \ker L = \dim V$$

the image of V under $v \longrightarrow \lambda_v$ in V is that of V. Since proper subspaces have strictly smaller dimension it must be that $L(V) = V^*$. ///

Let V be a finite-dimensional vector space with non-degenerate form $\langle,\rangle$, and W a subspace. Define the **orthogonal complement**

$$W^\perp = \{\lambda \in V^* : \lambda(w) = 0, \text{ for all } w \in W\}$$

[3.0.5] **Corollary:** Let V be a finite-dimensional vector space with a non-degenerate form $\langle,\rangle$, and W a subspace. Under the isomorphism $L : v \longrightarrow \lambda_v$ of V to its dual,

$$L\left(\{v \in V : \langle v, w\rangle = 0 \text{ for all } w \in W\}\right) = W^\perp$$

Proof: Suppose that $L(v) \in W^\perp$. Thus, $\lambda_v(w) = 0$ for all $w \in W$. That is, $\langle v, w\rangle = 0$ for all $w \in W$. On the other hand, suppose that $\langle v, w\rangle = 0$ for all $w \in W$. Then $\lambda_v(w) = 0$ for all $w \in W$, so $\lambda_v \in W^\perp$. ///

[3.0.6] **Corollary:** Now suppose that $\langle,\rangle$ is *symmetric*, meaning that $\langle v, w\rangle = \langle w, v\rangle$ for all $v, w \in V$. Redefine

$$W^\perp = \{v \in V : \langle v, w\rangle = 0 \text{ for all } w \in W\}$$

Then

$$\dim W + \dim W^\perp = \dim V$$

and

$$W^{\perp\perp} = W$$

Proof: With our original definition of $W^\perp_{\text{orig}}$ as

$$W^\perp_{\text{orig}} = \{\lambda \in V^* : \lambda(w) = 0 \text{ for all } w \in W\}$$

we had proven

$$\dim W + \dim W^\perp_{\text{orig}} = \dim V$$

We just showed that $L(W^\perp) = W^\perp_{\text{orig}}$, and since the map $L : V \longrightarrow V^*$ by $v \longrightarrow \lambda_v$ is an isomorphism

$$\dim W^\perp = \dim W^\perp_{\text{orig}}$$

Thus,

$$\dim W + \dim W^\perp = \dim V$$

as claimed.

Next, we claim that $W \subset W^{\perp\perp}$. Indeed, for $w \in W$ it is certainly true that for $v \in W^\perp$

$$\langle v, w \rangle = \langle v, w \rangle = 0$$

That is, we see easily that $W \subset W^{\perp\perp}$. On the other hand, from

$$\dim W + \dim W^\perp = \dim V$$

and

$$\dim W^\perp + \dim W^{\perp\perp} = \dim V$$

we see that $\dim W^{\perp\perp} = \dim W$. Since W is a subspace of $W^{\perp\perp}$ with the same dimension, the two must be equal (from our earlier discussion). ///

[3.0.7] **Remark:** When a non-degenerate bilinear form on V is *not* symmetric, there are two different versions of $W^\perp$, depending upon which argument in $\langle , \rangle$ is used:

$$W^{\perp,\mathrm{rt}} = \{v \in V : \langle v, w \rangle = 0, \quad \text{for all } w \in W\}$$
$$W^{\perp,\mathrm{lft}} = \{v \in V : \langle w, v \rangle = 0, \quad \text{for all } w \in W\}$$

And then there are two correct statements about $W^{\perp\perp}$, namely

$$\left(W^{\perp,\mathrm{rt}}\right)^{\perp,\mathrm{lft}} = W$$

$$\left(W^{\perp,\mathrm{lft}}\right)^{\perp,\mathrm{rt}} = W$$

These are proven in the same way as the last corollary, but with more attention to the lack of symmetry in the bilinear form. In fact, to more scrupulously consider possible asymmetry of the form, we proceed as follows.

For many purposes we can consider bilinear maps [11] (that is, k-valued maps linear in each argument)

$$\langle , \rangle : V \times W \longrightarrow k$$

where V and W are vectorspaces over the field k. [12]

The most common instance of such a pairing is that of a vector space and its dual

$$\langle , \rangle : V \times V^* \longrightarrow k$$

by

$$\langle v, \lambda \rangle = \lambda(v)$$

This notation and viewpoint helps to emphasize the near-symmetry [13] of the relationship between V and V^*.

[11] Also called bilinear *forms*, or bilinear *pairings*, or simply *pairings*.

[12] Note that now the situation is unsymmetrical, insofar as the first and second arguments to $\langle , \rangle$ are from different spaces, so that there is no obvious sense to any property of *symmetry*.

[13] The second dual V^{**} is naturally isomorphic to V if and only if $\dim V < \infty$.

Rather than simply *assume* non-degeneracy conditions, let us give ourselves a language to talk about such issues. Much as earlier, define

$$W^\perp = \{v \in V : \langle v, w\rangle = 0 \text{ for all } w \in W\}$$

$$V^\perp = \{w \in W : \langle v, w\rangle = 0 \text{ for all } v \in V\}$$

Then we have

[3.0.8] **Proposition:** A bilinear form $\langle,\rangle : V \times W \longrightarrow k$ induces a bilinear form, still denoted $\langle,\rangle$,

$$\langle,\rangle : V/W^\perp \times W/V^\perp \longrightarrow k$$

defined in the natural manner by

$$\langle v + W^\perp, w + V^\perp\rangle = \langle v, w\rangle$$

for any representatives v, w for the cosets. This form is *non-degenerate* in the sense that, on the quotient, given $x \in V/W^\perp$, there is $y \in W/V^\perp$ such that $\langle x, y\rangle \neq 0$, and symmetrically.

Proof: The first point is that the bilinear form on the quotients is *well-defined*, which is immediate from the definition of $W^\perp$ and $V^\perp$. Likewise, the non-degeneracy follows from the definition: given $x = v + W^\perp$ in $V/W^\perp$, take $w \in W$ such that $\langle v, w\rangle \neq 0$, and let $y = w + V^\perp$. ///

[3.0.9] **Remark:** The pairing of a vector space V and its dual is non-degenerate, even if the vector space is infinite-dimensional.

In fact, the pairing of (finite-dimensional) V and V^* is the universal example of a non-degenerate pairing:

[3.0.10] **Proposition:** For finite-dimensional V and W, a non-degenerate pairing

$$\langle,\rangle : V \times W \longrightarrow k$$

gives natural isomorphisms

$$V \xrightarrow{\approx} W^*$$

$$W \xrightarrow{\approx} V^*$$

via

$$v \longrightarrow \lambda_v \quad \text{where} \quad \lambda_v(w) = \langle v, w\rangle$$

$$w \longrightarrow \lambda_w \quad \text{where} \quad \lambda_w(v) = \langle v, w\rangle$$

Proof: The indicated maps are easily seen to be linear, with trivial kernels because the pairing is non-degenerate. The dimensions match, so these maps are isomorphisms. ///

---

4. *Worked examples*

[25.1] Let k be a field, and V a finite-dimensional k-vectorspace. Let Λ be a subset of the dual space V^*, with $|\Lambda| < \dim V$. Show that the **homogeneous system of equations**

$$\lambda(v) = 0 \ \ (\text{for all } \lambda \in \Lambda)$$

has a non-trivial (that is, non-zero) solution $v \in V$ (meeting all these conditions).

The dimension of the span W of Λ is strictly less than $\dim V^*$, which we've proven is $\dim V^* = \dim V$. We may also identify $V \approx V^{**}$ via the natural isomorphism. With that identification, we may say that the set of solutions is $W^\perp$, and

$$\dim(W^\perp) + \dim W = \dim V^* = \dim V$$

Thus, $\dim W^\perp > 0$, so there are non-zero solutions. ///

[25.2] Let k be a field, and V a finite-dimensional k-vectorspace. Let Λ be a *linearly independent* subset of the dual space V^*. Let $\lambda \longrightarrow a_\lambda$ be a set map $\Lambda \longrightarrow k$. Show that an **inhomogeneous system of equations**

$$\lambda(v) = a_\lambda \ \text{ (for all } \lambda \in \Lambda)$$

has a solution $v \in V$ (meeting all these conditions).

Let $m = |\Lambda|$, $\Lambda = \{\lambda_1, \ldots, \lambda_m\}$. One way to use the linear independence of the functionals in Λ is to extend Λ to a basis $\lambda_1, \ldots, \lambda_n$ for V^*, and let $e_1, \ldots, e_n \in V^{**}$ be the corresponding dual basis for V^{**}. Then let $v_1, \ldots, v_n$ be the images of the e_i in V under the natural isomorphism $V^{**} \approx V$. (This achieves the effect of making the λ_i be a dual basis to the v_i. We had only literally proven that one can go from a basis of a vector space to a dual basis of its dual, and not the reverse.) Then

$$v = \sum_{1 \le i \le m} a_{\lambda_i} \cdot v_i$$

is a solution to the indicated set of equations, since

$$\lambda_j(v) = \sum_{1 \le i \le m} a_{\lambda_i} \cdot \lambda_j(v_i) = a_{\lambda_j}$$

for all indices $j \le m$. ///

[25.3] Let T be a k-linear endomorphism of a finite-dimensional k-vectorspace V. For an eigenvalue λ of T, let V_λ be the generalized λ-eigenspace

$$V_\lambda = \{v \in V : (T - \lambda)^n v = 0 \text{ for some } 1 \le n \in \mathbb{Z}\}$$

Show that the projector P of V to V_λ (commuting with T) lies inside $k[T]$.

First we do this assuming that the minimal polynomial of T factors into linear factors in $k[x]$.

Let $f(x)$ be the minimal polynomial of T, and let $f_\lambda(x) = f(x)/(x-\lambda)^e$ where $(x-\lambda)^e$ is the precise power of $(x - \lambda)$ dividing $f(x)$. Then the collection of all $f_\lambda(x)$'s has *gcd* 1, so there are $a_\lambda(x) \in k[x]$ such that

$$1 = \sum_\lambda a_\lambda(x)\, f_\lambda(x)$$

We claim that $E_\lambda = a_\lambda(T) f_\lambda(T)$ is a projector to the generalized λ-eigenspace V_λ. Indeed, for $v \in V_\lambda$,

$$v = 1_V \cdot v = \sum_\mu a_\mu(T) f_\mu(T) \cdot v = \sum_\mu a_\mu(T) f_\mu(T) \cdot v = a_\lambda(T) f_\lambda(T) \cdot v$$

since $(x-\lambda)^e$ divides $f_\mu(x)$ for $\mu \ne \lambda$, and $(T-\lambda)^e v = 0$. That is, it acts as the identity on V_λ. And

$$(T-\lambda)^e \circ E_\lambda = a_\lambda(T)\, f(T) = 0 \in \mathrm{End}_k(V)$$

so the image of E_λ is inside V_λ. Since E_λ is the identity on V_λ, it must be that the image of E_λ is *exactly* V_λ. For $\mu \ne \lambda$, since $f(x) | f_\mu(x) f_\lambda(x)$, $E_\mu E_\lambda = 0$, so these idempotents are *mutually orthogonal*. Then

$$(a_\lambda(T) f_\lambda(T))^2 = (a_\lambda(T) f_\lambda(T)) \cdot (1 - \sum_{\mu \ne \lambda} a_\mu(T) f_\mu(T)) = a_\lambda(T) f_\lambda(T) - 0$$

That is, $E_\lambda^2 = E_\lambda$, so E_λ is *a* projector to V_λ.

The mutual orthogonality of the idempotents will yield the fact that V is the direct sum of all the generalized eigenspaces of T. Indeed, for any $v \in V$,

$$v = 1 \cdot v = (\sum_\lambda E_\lambda)\, v = \sum_\lambda (E_\lambda v)$$

and $E_\lambda v \in V_\lambda$. Thus,

$$\sum_\lambda V_\lambda = V$$

To check that the sum is (unsurprisingly) direct, let $v_\lambda \in V_\lambda$, and suppose

$$\sum_\lambda v_\lambda = 0$$

Then $v_\lambda = E_\lambda v_\lambda$, for all λ. Then apply E_μ and invoke the orthogonality of the idempotents to obtain

$$v_\mu = 0$$

This proves the linear independence, and that the sum is direct.

To prove *uniqueness* of a projector E to V_λ commuting with T, note that any operator S commuting with T necessarily stabilizes all the generalized eigenspaces of T, since for $v \in V_\mu$

$$(T-\lambda)^e\, Sv = S\,(T-\lambda)^e v = S \cdot 0 = 0$$

Thus, E stabilizes all the V_μs. Since V is the direct sum of the V_μ and E maps V to V_λ, it must be that E is 0 on V_μ for $\mu \neq \lambda$. Thus,

$$E = 1 \cdot E_\lambda + \sum_{\mu \neq \lambda} 0 \cdot E_\mu = E_\lambda$$

That is, there is just one projector to V_λ that also commutes with T. *This finishes things under the assumption that $f(x)$ factors into linear factors in $k[x]$.*

The more general situation is similar. More generally, for a monic irreducible $P(x)$ in $k[x]$ dividing $f(x)$, with $P(x)^e$ the precise power of $P(x)$ dividing $f(x)$, let

$$f_P(x) = f(x)/P(x)^e$$

Then these f_P have *gcd* 1, so there are $a_P(x)$ in $k[x]$ such that

$$1 = \sum_P a_P(x) \cdot f_P(x)$$

Let $E_P = a_P(T) f_P(T)$. Since $f(x)$ divides $f_P(x) \cdot f_Q(x)$ for distinct irreducibles P, Q, we have $E_P \circ E_Q = 0$ for $P \neq Q$. And

$$E_P^2 = E_P(1 - \sum_{Q \neq P} E_Q) = E_P$$

so (as in the simpler version) the E_P's are mutually orthogonal idempotents. And, similarly, V is the direct sum of the subspaces

$$V_P = E_P \cdot V$$

We can also characterize V_P as the kernel of $P^e(T)$ on V, where $P^e(x)$ is the power of $P(x)$ dividing $f(x)$. If $P(x) = (x - \lambda)$, then V_P is the generalized λ-eigenspace, and E_P is the projector to it.

If E were another projector to V_λ commuting with T, then E stabilizes V_P for all irreducibles P dividing the minimal polynomial f of T, and E is 0 on V_Q for $Q \neq (x-\lambda)$, and E is 1 on V_λ. That is,

$$E = 1 \cdot E_{x-\lambda} + \sum_{Q \neq x-\lambda} 0 \cdot E_Q = E_P$$

This proves the uniqueness even in general. ///

[25.4] Let T be a matrix in Jordan normal form with entries in a field k. Let T_{ss} be the matrix obtained by converting all the off-diagonal 1's to 0's, making T diagonal. Show that T_{ss} is in $k[T]$.

This implicitly demands that the minimal polynomial of T factors into linear factors in $k[x]$.

Continuing as in the previous example, let $E_\lambda \in k[T]$ be the projector to the generalized λ-eigenspace V_λ, and keep in mind that we have shown that V is the direct sum of the generalized eigenspaces, equivalent, that $\sum_\lambda E_\lambda = 1$. By definition, the operator T_{ss} is the scalar operator λ on V_λ. Then

$$T_{ss} = \sum_\lambda \lambda \cdot E_\lambda \in k[T]$$

since (from the previous example) each E_λ is in $k[T]$. ///

[25.5] Let $M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$ be a matrix in a block decomposition, where A is m-by-m and D is n-by-n. Show that

$$\det M = \det A \cdot \det D$$

One way to prove this is to use the formula for the determinant of an N-by-N matrix

$$\det C = \sum_{\pi \in S_N} \sigma(\pi)\, a_{\pi(1),1} \ldots a_{\pi(N),N}$$

where c_{ij} is the $(i,j)^{th}$ entry of C, π is summed over the symmetric group S_N, and σ is the sign homomorphism. Applying this to the matrix M,

$$\det M = \sum_{\pi \in S_{m+n}} \sigma(\pi)\, M_{\pi(1),1} \ldots M_{\pi(m+n),m+n}$$

where M_{ij} is the $(i,j)^{th}$ entry. Since the entries M_{ij} with $1 \le j \le m$ and $m < i \le m+n$ are all 0, we should only sum over π with the property that

$$\pi(j) \le m \quad \text{for} \quad 1 \le j \le m$$

That is, π stabilizes the subset $\{1, \ldots, m\}$ of the indexing set. Since π is a bijection of the index set, necessarily such π stabilizes $\{m+1, m+2, \ldots, m+n\}$, also. Conversely, each pair (π_1, π_2) of permutation π_1 of the first m indices and π_2 of the last n indices gives a permutation of the whole set of indices.

Let X be the set of the permutations $\pi \in S_{m+n}$ that stabilize $\{1, \ldots, m\}$. For each $\pi \in X$, let π_1 be the restriction of π to $\{1, \ldots, m\}$, and let π_2 be the restriction to $\{m+1, \ldots, m+n\}$. And, in fact, if we plan to index the entries of the block D in the usual way, we'd better be able to think of π_2 as a permutation of $\{1, \ldots, n\}$, also. Note that $\sigma(\pi) = \sigma(\pi_1)\sigma(\pi_2)$. Then

$$\det M = \sum_{\pi \in X} \sigma(\pi)\, M_{\pi(1),1} \ldots M_{\pi(m+n),m+n}$$

$$= \sum_{\pi \in X} \sigma(\pi)\,(M_{\pi(1),1} \ldots M_{\pi(m),m}) \cdot (M_{\pi(m+1),m+1} \ldots M_{\pi(m+n),m+n})$$

$$= \left(\sum_{\pi_1 \in S_m} \sigma(\pi_1)\, M_{\pi_1(1),1} \ldots M_{\pi_1(m),m} \right) \cdot \left(\sum_{\pi_2 \in S_n} \sigma(\pi_2)(M_{\pi_2(m+1),m+1} \ldots M_{\pi_2(m+n),m+n} \right)$$

$$= \left(\sum_{\pi_1 \in S_m} \sigma(\pi_1)\, A_{\pi_1(1),1} \ldots A_{\pi_1(m),m} \right) \cdot \left(\sum_{\pi_2 \in S_n} \sigma(\pi_2) D_{\pi_2(1),1} \ldots D_{\pi_2(n),n} \right) = \det A \cdot \det D$$

where in the last part we have mapped $\{m+1, \ldots, m+n\}$ bijectively by $\ell \longrightarrow \ell - m$. ///

[25.6] The so-called *Kronecker product*[14] of an m-by-m matrix A and an n-by-n matrix B is

$$A \otimes B = \begin{pmatrix} A_{11} \cdot B & A_{12} \cdot B & \ldots & A_{1m} \cdot B \\ A_{21} \cdot B & A_{22} \cdot B & \ldots & A_{2m} \cdot B \\ & \vdots & & \\ A_{m1} \cdot B & A_{m2} \cdot B & \ldots & A_{mm} \cdot B \end{pmatrix}$$

where, as it may appear, the matrix B is inserted as n-by-n blocks, multiplied by the respective entries A_{ij} of A. Prove that

$$\det(A \otimes B) = (\det A)^n \cdot (\det B)^m$$

at least for $m = n = 2$.

If no entry of the first row of A is non-zero, then both sides of the desired equality are 0, and we're done. So suppose some entry A_{1i} of the first row of A is non-zero. If $i \neq 1$, then for $\ell = 1, \ldots, n$ interchange the ℓ^{th} and $(i-1)n + \ell^{th}$ columns of $A \otimes B$, thus multiplying the determinant by $(-1)^n$. This is compatible with the formula, so we'll assume that $A_{11} \neq 0$ to do an induction on m.

We will manipulate n-by-n *blocks* of scalar multiples of B rather than actual scalars.

Thus, assuming that $A_{11} \neq 0$, we want to subtract multiples of the left column of n-by-n blocks from the blocks further to the right, to make the top n-by-n blocks all 0 (apart from the leftmost block, $A_{11}B$). In terms of manipulations of columns, for $\ell = 1, \ldots, n$ and $j = 2, 3, \ldots, m$ subtract A_{1j}/A_{11} times the ℓ^{th} column of $A \otimes B$ from the $((j-1)n + \ell)^{th}$. Since for $1 \leq \ell \leq n$ the ℓ^{th} column of $A \otimes B$ is A_{11} times the ℓ^{th} column of B, and the $((j-1)n + \ell)^{th}$ column of $A \otimes B$ is A_{1j} times the ℓ^{th} column of B, this has the desired effect of killing off the n-by-n blocks along the top of $A \otimes B$ except for the leftmost block. And the $(i,j)^{th}$ n-by-n block of $A \otimes B$ has become $(A_{ij} - A_{1j}A_{i1}/A_{11}) \cdot B$. Let

$$A'_{ij} = A_{ij} - A_{1j}A_{i1}/A_{11}$$

and let D be the $(m-1)$-by-$(m-1)$ matrix with $(i,j)^{th}$ entry $D_{ij} = A'_{(i-1),(j-1)}$. Thus, the manipulation so far gives

$$\det(A \otimes B) = \det \begin{pmatrix} A_{11}B & 0 \\ * & D \otimes B \end{pmatrix}$$

By the previous example (or its tranpose)

$$\det \begin{pmatrix} A_{11}B & 0 \\ * & D \otimes B \end{pmatrix} = \det(A_{11}B) \cdot \det(D \otimes B) = A_{11}^n \det B \cdot \det(D \otimes B)$$

by the multilinearity of det.

[14] As we will see shortly, this is really a **tensor product**, and we will treat this question more sensibly.

And, at the same time subtracting A_{1j}/A_{11} times the first column of A from the j^{th} column of A for $2 \le j \le m$ does not change the determinant, and the new matrix is

$$\begin{pmatrix} A_{11} & 0 \\ * & D \end{pmatrix}$$

Also by the previous example,

$$\det A = \det \begin{pmatrix} A_{11} & 0 \\ * & D \end{pmatrix} = A_1 1 \cdot \det D$$

Thus, putting the two computations together,

$$\det(A \otimes B) = A_{11}^n \det B \cdot \det(D \otimes B) = A_{11}^n \det B \cdot (\det D)^n (\det B)^{m-1}$$
$$= (A_{11} \det D)^n \det B \cdot (\det B)^{m-1} = (\det A)^n (\det B)^m$$

as claimed.

Another approach to this is to observe that, in these terms, $A \otimes B$ is

$$\begin{pmatrix} A_{11} & 0 & \ldots & 0 & & A_{1m} & 0 & \ldots & 0 \\ 0 & A_{11} & & & & 0 & A_{1m} & & \\ \vdots & & \ddots & & \cdots & \vdots & & \ddots & \\ 0 & & & A_{11} & & 0 & & & A_{1m} \\ & \vdots & & & & & \vdots & & \\ A_{m1} & 0 & \ldots & 0 & & A_{mm} & 0 & \ldots & 0 \\ 0 & A_{m1} & & & & 0 & A_{mm} & & \\ \vdots & & \ddots & & \cdots & \vdots & & \ddots & \\ 0 & & & A_{m1} & & 0 & & & A_{mm} \end{pmatrix} \begin{pmatrix} B & 0 & \ldots & 0 \\ 0 & B & & \\ \vdots & & \ddots & \\ 0 & & & B \end{pmatrix}$$

where there are m copies of B on the diagonal. By suitable permutations of rows and columns (with an interchange of rows for each interchange of columns, thus giving no net change of sign), the matrix containing the A_{ij}s becomes

$$\begin{pmatrix} A & 0 & \ldots & 0 \\ 0 & A & & \\ \vdots & & \ddots & \\ 0 & & & A \end{pmatrix}$$

with n copies of A on the diagonal. Thus,

$$\det(A \otimes B) = \det \begin{pmatrix} A & 0 & \ldots & 0 \\ 0 & A & & \\ \vdots & & \ddots & \\ 0 & & & A \end{pmatrix} \cdot \det \begin{pmatrix} B & 0 & \ldots & 0 \\ 0 & B & & \\ \vdots & & \ddots & \\ 0 & & & B \end{pmatrix} = (\det A)^n \cdot (\det B)^m$$

This might be more attractive than the first argument, depending on one's tastes. ///

Exercises

25.[4.0.1] Let T be a hermitian operator on a finite-dimensional complex vector space V with a positive-definite inner product $\langle , \rangle$. Let P be an orthogonal projector to the λ-eigenspace V_λ of T. (This means that P is the identity on V_λ and is 0 on the orthogonal complement $V_\lambda^\perp$ of V_λ.) Show that $P \in \mathbb{C}[T]$.

25.[4.0.2] Let T be a diagonalizable operator on a finite-dimensional vector space V over a field k. Show that there is a *unique* projector P to the λ-eigenspace V_λ of T such that $TP = PT$.

25.[4.0.3] Let k be a field, and V, W finite-dimensional vector spaces over k. Let S be a k-linear endomorphism of V, and T a k-linear endomorphism of W. Let $S \oplus T$ be the k-linear endomorphism of $V \oplus W$ defined by

$$(S \oplus T)(v \oplus w) \;=\; S(v) \oplus T(w) \qquad\qquad (\text{for } v \in V \text{ and } w \in W)$$

Show that the minimal polynomial of $S \oplus T$ is the least common multiple of the minimal polynomials of S and T.

25.[4.0.4] Let T be an n-by-n matrix with entries in a commutative ring R, with non-zero entries only *above* the diagonal. Show that $T^n = 0$.

25.[4.0.5] Let T be an endomorphism of a finite-dimensional vector space V over a field k. Suppose that T is *nilpotent*, that is, that $T^n = 0$ for some positive integer n. Show that $\operatorname{tr} T = 0$.

25.[4.0.6] Let k be a field of characteristic 0, and T a k-linear endomorphism of an n-dimensional vector space V over k. Show that T is nilpotent if and only if $\operatorname{trace}(T^i) = 0$ for $1 \le i \le n$.

25.[4.0.7] Fix a field k of characteristic not 2, and let $K = k(\sqrt{D})$ where D is a non-square in k. Let σ be the non-trivial automorphism of K over k. Let $\Delta \in k^\times$. Let A be the k-subalgebra of 2-by-2 matrices over K generated by

$$\begin{pmatrix} 0 & 1 \\ \Delta & 0 \end{pmatrix} \quad \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^\sigma \end{pmatrix}$$

where α ranges over K. Find a condition relating D and Δ necessary and sufficient for A to be a division algebra.

25.[4.0.8] A *Lie algebra* (named after the mathematician Sophus Lie) over a field k of characteristic 0 is a k-vectorspace with a k-bilinear map $[,]$ (the *Lie bracket*) such that $[x, y] = -[y, x]$, and satisfying the *Jacobi identity*

$$[[x, y], z] = [x, [y, z]] - [y, [x, z]]$$

Let A be an (associative) k-algebra. Show that A can be made into a Lie algebra by defining $[x, y] = xy - yx$.

25.[4.0.9] Let $\mathfrak{g}$ be a Lie algebra over a field k. Let A be the associative algebra of k-vectorspace endomorphisms of $\mathfrak{g}$. The *adjoint* action of $\mathfrak{g}$ on itself is defined by

$$(\mathrm{ad}x)(y) \;=\; [x, y]$$

Show that the map $\mathfrak{g} \longrightarrow \mathrm{Aut}_k G$ defined by $x \longrightarrow \mathrm{ad}x$ is a *Lie homomorphism*, meaning that

$$[\mathrm{ad}x, \mathrm{ad}y] \;=\; \mathrm{ad}[x, y]$$

(The latter property is the *Jacobi identity*.)