Chapter III. Dual Spaces and Duality.

III.1 Definitions and Examples.

The **linear functionals** (also known as **dual vectors**) on a vector space V over $\mathbb{K}$ are the linear maps $\ell : V \rightarrow \mathbb{K}$. We denote the space of functionals by V^*, or equivalently $\mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$. It becomes a vector space when we impose the operations

1. Addition: $\big(\ell_1 + \ell_2\big)(v) = \ell_1(v) + \ell_2(v)$ for all $v \in V$
2. Scaling: $\big(\lambda\cdot\ell\big)(v) = \lambda\cdot\ell(v)$ for $\lambda \in \mathbb{K}, v \in V$

The zero element in V^* is the **zero functional** $\ell(v) = 0_{\mathbb{K}}$ for all $v \in V$, for which $\ker(\ell) = V$ and $\mathrm{range}(\ell) = \{0_{\mathbb{K}}\}$.

Notation: We will often employ "bracket" notation in discussing functionals, writing

$$\langle \ell, v\rangle \quad \text{instead of} \quad \ell(v)$$

This notation combines inputs ℓ, v to create a map $V^* \times V \rightarrow \mathbb{K}$ that is linear in each entry when the other entry is held fixed. In bracket notation both inputs play equal roles, and either one can be held fixed while the other varies. As we shall see this has many advantages. $\square$

We begin with an example that is central in understanding what dual vectors are and what they do.

1.1. Example. Let V be a finite dimensional space and $\mathfrak{X} = \{e_1, \ldots, e_n\}$ an ordered basis. Every $v \in V$ has a unique expansion

$$v = \sum_{i=1}^{n} c_i e_i \qquad (c_i \in \mathbb{K})$$

For each $1 \leq i \leq n$ the map $e_i^* : V \rightarrow \mathbb{K}$ that reads off the i^{th} coefficient

$$\langle e_i^*, v\rangle = c_i$$

is a linear functional in V^*. We will soon see that the set of functionals $\mathfrak{X}^* = \{e_1^*, \ldots, e_n^*\}$ is a basis for the dual space V^*, called the **dual basis** determined by $\mathfrak{X}$, from which it follows that the dual space is finite dimensional with $\dim(V^*) = \dim(V) = n$. $\square$

The following examples give some idea of the ubiquity of dual spaces in linear algebra.

1.2. Example. For $V = \mathbb{K}[x]$ an element $a \in \mathbb{K}$ determines an "evaluation functional" $\epsilon_a \in V^*$:

$$\langle \epsilon_a, f\rangle = \sum_{k=0}^{n} c_k a^k \qquad \text{if} \qquad f = \sum_{k=0}^{n} c_k x^k$$

These do not by themselves form a vector subspace of V^* because $\langle \epsilon_a - \epsilon_b, f\rangle = f(a) - f(b)$ cannot always be written as $f(c)$ for some $c \in \mathbb{K}$.

More generally, if $V = \mathcal{C}[a, b]$ is the space of *continuous* complex valued functions on the interval $X = [a, b] \subseteq \mathbb{R}$ we can define evaluation functionals $\langle \epsilon_s, f\rangle = f(s)$ for

$a \le s \le b$, but many element in V^* are of a quite different nature. Two examples:

$$\text{(i)} \qquad I(f) = \int_a^b f(t)\,dt \qquad \text{(Riemann integral of } f\text{)}$$

$$\text{(ii)} \qquad I^x(f) = \int_a^x f(t)\,dt \quad \text{(for any endpoint } a \le x \le b\text{)}$$

For another example, consider the space $V = \mathcal{C}^{(1)}(a,b)$ of real-valued functions on an interval $(a,b) \subseteq \mathbb{R}$ that have continuous first derivative $df/dx(s)$. We can define the usual evaluation functionals $\epsilon_s \in V^*$, but since differentiation is a linear operator on $\mathcal{C}^{(1)}(a,b)$ there are also functionals ℓ_s involving derivatives, such as

$$\ell_s : f \to \frac{df}{dx}(s) \qquad \text{for } a < s < b \ ,$$

or even linear combinations such as $\tilde{\ell}_s(f) = f(s) + \frac{df}{dx}(s)$. $\square$

1.3. Example. Suppose V is finite dimensional and that $l \in V^*$ is not the zero functional. The kernel $E = \ker(\ell) = \{v \in V : \langle \ell, v\rangle = 0\}$ is a "*hyperplane*" in V – a vector subspace of dimension $n-1$ where $n = \dim(V)$.

Proof: By the dimension formula,

$$\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}\big(\ker(\ell)\big) + \dim_{\mathbb{K}}\big(\mathrm{range}(\ell)\big)$$

But if $\ell \neq 0$, say $\langle \ell, v_0\rangle \neq 0$, then $\langle \ell, \mathbb{K}v_0\rangle = \mathbb{K}$, so $\mathrm{range}(\ell) = \mathbb{K}$ has dimension 1. $\square$

1.4. Example. On $\mathbb{R}^n$ we have the standard Euclidean inner product

$$(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^n x_k y_k \qquad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \ ,$$

familiar from Calculus, but this is just a special case of the standard inner product on complex n-dimensional coordinate space $\mathbb{C}^n$,

$$(\mathbf{z}, \mathbf{w}) = \sum_{k=1}^n z_k \overline{w_k} \qquad \text{for complex } n\text{-tuples } \mathbf{z}, \mathbf{w} \text{ in } \mathbb{C}^n \ , \tag{23}$$

where $\overline{z} = x - iy$ is the complex conjugate of $z = x + iy$. We will focus on the complex case, because everything said here applies verbatim to the real case if you interpret "complex conjugation" to mean $\overline{x} = x$ for real numbers.

In either case, imposing an inner product on coordinate space $V = \mathbb{K}^n$ allows us to construct $\mathbb{K}$-linear functionals $\ell_{\mathbf{y}} \in V^*$ associated with individual vectors $\mathbf{y} \in V = \mathbb{K}^n$, by defining

$$\langle \ell_{\mathbf{y}}, \mathbf{x}\rangle = (\mathbf{x}, \mathbf{y}) \qquad \text{for any } \mathbf{x} \in V$$

In this setting the right hand vector $\mathbf{y}$ is fixed, and acts on the left-hand entry to produce a scalar in $\mathbb{K}$. (Think of $\mathbf{y}$ as the "actor" and $\mathbf{x}$ as the "actee" – the vector that gets acted upon.)

The functional $\ell_{\mathbf{y}}$ is $\mathbb{K}$-linear because the inner product is linear in its first entry when the second entry $\mathbf{y}$ is held fixed, hence $\ell_{\mathbf{y}}$ is a dual vector in V^*. Note carefully the placement of the "actee" on the left side of the inner product; the inner product on a vector space over $\mathbb{K} = \mathbb{C}$ is a *conjugate-linear* function of the right hand entry.

$$(\mathbf{z}, \lambda \cdot \mathbf{w}) = \overline{\lambda} \cdot (\mathbf{z}, \mathbf{w}) \quad \text{ while } \quad (\lambda \cdot \mathbf{z}, \mathbf{w}) = \lambda \cdot (\mathbf{z}, \mathbf{w})$$

for $\lambda \in \mathbb{K}$. Placing the "actee" on the right would not produce a $\mathbb{C}$-*linear* operation on input vectors. (When $\mathbb{K} = \mathbb{R}$, complex conjugation doesn't do anything, and "conjugate-linear" is the same as "linear.")

The special case $\mathbb{K}^n = \mathbb{R}^n$ is of course important in geometry. The inner product on $\mathbb{R}^n$ and the functionals $\ell_{\mathbf{y}}$ then have explicit geometric interpretations:

$$\begin{aligned}
(\mathbf{x}, \mathbf{y}) &= \langle \ell_{\mathbf{y}}, \mathbf{x} \rangle = \|\mathbf{x}\| \cdot \|\mathbf{y}\| \cdot \cos(\theta) \\
&= \|\mathbf{x}\| \cdot \big(\|\mathbf{y}\| \cdot \cos\,\theta \big) \\
&= \|\mathbf{x}\| \cdot \left(\begin{array}{c} \text{orthogonally projected length of } \mathbf{y} \\ \text{on the 1-dimensional subspace } \mathbb{R}{\cdot}\mathbf{x} \end{array} \right) \quad ,
\end{aligned}$$

where

$$\|\mathbf{x}\| = (\mathbf{x}, \mathbf{x})^{1/2} = \Big(\sum_{k=1}^{n} |x_k|^2 \Big)^{1/2}$$

is the Euclidean length of vector $\mathbf{x} \in \mathbb{R}^n$. The angle $\theta = \theta(\mathbf{x}, \mathbf{y})$ is the angle in radians between $\mathbf{x}$ and $\mathbf{y}$, measured in the plane (two-dimensional subspace) spanned by $\mathbf{x}$ and $\mathbf{y}$ as shown in Figure 3.1. Notice that $\mathbf{x}$ and $\mathbf{y}$ are perpendicular if $(\mathbf{x}, \mathbf{y}) = 0$, so $\cos(\theta) = 0$.

Note: While the real inner product is natural in geometry, in physics the complex inner product is the notion of choice (in electrical engineering, quantum mechanics, etc, etc). *But beware*: physicists employ a convention opposite to ours. For them an inner product is linear in the *right-hand* entry and conjugate linear on the *left.* That can be confusing if you are not forwarned. $\square$

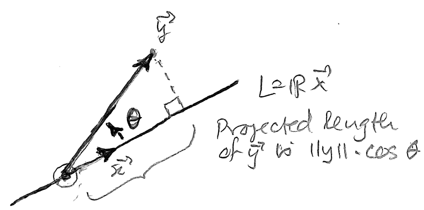

Figure 3.1. Geometric interpretation of the standard inner product $(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}\|\,\|\mathbf{y}\| \cdot \cos(\theta(\mathbf{x}, \mathbf{y}))$ in $\mathbb{R}^n$. The projected length of a vector $\mathbf{y}$ onto the line $L = \mathbb{R}\mathbf{x}$ is $\|\mathbf{y}\| {\cdot} \cos(\theta)$. The angle $\theta(\mathbf{x}, \mathbf{y})$ is measured within the two-dimensional subspace $M = \mathbb{R}\text{-span}\{\mathbf{x}, \mathbf{y}\}$. Vectors are *orthogonal* when $(\mathbf{x}, \mathbf{y}) = 0$, so $\cos\,\theta = 0$. The zero vector is orthogonal to everybody.

1.5. Example. In $V = \mathbb{R}^3$ with the standard inner product $(\mathbf{x}, \mathbf{y}) = \sum_i x_i y_i$, fix a vector $\mathbf{u} \neq \mathbf{0}$. The set of vectors $M = \{\mathbf{x} \in \mathbb{R}^3 : (\mathbf{x}, \mathbf{u}) = 0\}$ is the hyperplane of vectors orthogonal to $\mathbf{u}$ – see Figure 3.2. As an example, if $\mathbf{u} = (1, 0, 0) \in \mathbb{R}^3$ and $\ell_{\mathbf{u}}(x) = (\mathbf{x}, \mathbf{u})$ as in Example 1.4, this orthogonal hyperplane coincides with the kernel of $\ell_{\mathbf{u}}$:

$$M = \ker\big(\ell_{\mathbf{u}}\big) = \{(x_1, x_2, 0) : x_1, x_2 \in \mathbb{R}\} = \mathbb{R}\text{-span}\{\mathbf{e}_1, \mathbf{e}_2\} \quad \square\,.$$

1.6. Exercise. If $\mathbf{u} \neq \mathbf{0}$ in an inner product space of dimension n, explain why the orthogonal complement

$$M = \big(\mathbb{R}{\cdot}\mathbf{u}\big)^{\perp} = \{\mathbf{x} : (\mathbf{x}, \mathbf{u}) = 0\}$$

is a subspace of dimension $n - 1$.
Hint: Reread Example 1.3.

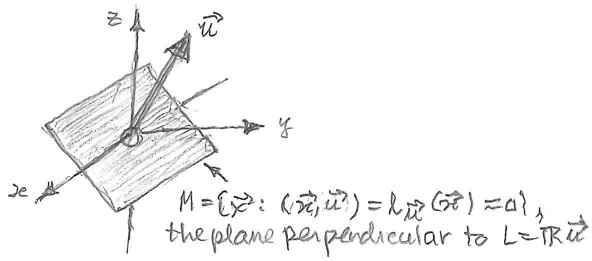

Figure 3.2. A nonzero vector u $\in \mathbb{R}^n$ determines a *hyperplane* $M = (\mathbb{R}\mathbf{u})^{\perp} = \{\mathbf{x} : (\mathbf{x}, \mathbf{y}) = 0\} = \ker(\ell_{\mathbf{u}})$, an $(n-1)$-dimensional subspace consisting of the vectors perpendicular to $\mathbf{u}$.

1.7. Example. Let $V = \mathcal{C}[0,1]$ be the ∞-dimensional space of all continuous complex-valued functions $f : [0,1] \rightarrow \mathbb{C}$. The **Fourier transform** of f is the function $f^\wedge : \mathbb{Z} \rightarrow \mathbb{C}$ defined by integrating $f(t)$ against the *complex trigonometric functions*

$$E_n(t) = e^{2\pi i n t} = \cos(2\pi t) + i \sin(2\pi t) \qquad (n \in \mathbb{Z})$$

on the real line. The n^{th} **Fourier coefficient** of $f(t)$ is the integral:

$$f^\wedge(n) = \int_0^1 f(t) e^{-2\pi i n t}\, dt \qquad (n \in \mathbb{Z})$$

(Note that the E_n are all periodic with period $\Delta t = 1$, so this integral is taken over the basic period $0 \leq t \leq 1$ common to them all.) If $f(t)$ is smooth and periodic with $f(t+1) = f(t)$ for all $t \in \mathbb{R}$, it can be synthesized as a superposition of the basic complex trigonometric functions E_n, with weights given by the Fourier coefficients:

$$f(t) = \sum_{n=-\infty}^{+\infty} f^\wedge(n) \cdot e^{2\pi i n t} = \sum_{n=-\infty}^{+\infty} f^\wedge(n) \cdot E_n(t)$$

The series converges pointwise on $\mathbb{R}$ if f is periodic and once continuously differentiable.

For each index $n \in \mathbb{Z}$ the map

$$f \in \mathcal{C}[0,1] \quad \xrightarrow{\phi_n} \quad f^\wedge(n) \in \mathbb{C}$$

is a linear functional in V^*. It is actually another example of a functional determined via an inner product as in Example 1.4. The standard inner product on $\mathcal{C}[0,1]$ is $(f,h) = \int_0^1 f(t)\overline{h(t)}\, dt$, and we have

$$\phi_n(f) = f^\wedge(n) = \int_0^1 f(t)\overline{E_n(t)}\, dt = \big(f, E_n\big)$$

for all $n \in \mathbb{Z}, f \in V$. So, ϕ_n is precisely the functional ℓ_{E_n} in Example 1.4. $\square$

III.2. Dual Bases in V*.

The dual space V^* of linear functionals can be viewed as the space of linear operators $\mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$. For arbitrary vector spaces V, W of dimension m, n we saw earlier in Lemma 4.4 of Chapter II that $\mathrm{Hom}_{\mathbb{K}}(V, W)$ is isomorphic to the space $\mathrm{M}(n \times m, \mathbb{K})$ of $n \times m$

matrices, which obviously has dimension $m \cdot n$. In the special case when $W = \mathbb{K}$ we get $\dim(V^*) = m = \dim(V)$.

This can also be seen by re-examining Example 1.1, which provides a natural way to construct a basis $\mathfrak{X}^* = \{e_1^*, \ldots, e_n^*\}$ in V^*, given an ordered basis $\mathfrak{X} = \{e_1, \ldots, e_n\}$ in V. The functional e_i^* reads the i^{th} coefficent in the unique expansion $v = \sum_i c_i e_i$ of a vector $v \in V$, so that

$$\Big\langle e_i^*, \sum_{k=1}^{n} c_k e_k \Big\rangle = c_i \qquad \text{for } 1 \leq i \leq n \tag{24}$$

As an immediate consequence, the linear functional $e_i^* : V \to \mathbb{K}$ is completely determined by the property

$$\langle\, e_i^*, e_j \,\rangle = \delta_{ij} \qquad \text{(the Kronecker delta symbol} = 1 \text{ if } i = j \text{ and } 0 \text{ otherwise)} \tag{25}$$

Identity (25) follows because $e_j = 0 \cdot e_1 + \ldots + 1 \cdot e_j + \ldots + 0 \cdot e_n$; we recover (24) by observing that

$$\Big\langle e_i^*, \sum_{k=1}^{n} c_k e_k \Big\rangle = \sum_{k=1}^{n} c_k \langle e_i^*, e_k \rangle = \sum_{k=1}^{n} c_k \delta_{ik} = c_i$$

as expected.

We now show that the vectors $e_1^*, \ldots, e_n^*$ form a basis in V^*, the **dual basis** to the original basis $\mathfrak{X}$ in V. This implies that $\dim(V^*) = \dim(V) = n$. Note, however, that to define the dual vectors e_i^* you must start with a basis in V; given a single vector "v" in V there is no way to define a dual vector "v^*" in V^*.

2.1. Theorem. *If V is finite dimensional and $\mathfrak{X}$ is a basis for V, the vectors $\mathfrak{X}^* = \{e_1^*, ..., e_n^*\}$ are a basis for V^*.*

Proof: Independence. If $\ell = \sum_{j=1}^{n} c_j e_j^*$ is the zero vector in V^* then $\langle \sum_j c_i e_j^*, v \rangle = 0$ for every $v \in V$, and in particular if $v = e_i$ we get

$$0 = \langle \ell, e_i \rangle = \sum_j c_j \langle e_j^*, e_i \rangle = \sum_j c_j \delta_{ji} = c_i$$

for $1 \leq i \leq n$, proving independence of the vectors e_i^*.

Spanning. If $\ell \in V^*$ and $c_i = \langle \ell, e_i \rangle$, we claim that ℓ is equal to $\ell' = \sum_{j=1}^{n} \langle \ell, e_j \rangle \cdot e_j^*$. It suffices to show that ℓ and ℓ' have the same values on the basis vectors $\{e_i\}$ in V, but that is obvious because

$$\begin{aligned} \langle \ell', e_i \rangle &= \Big\langle \sum_j \langle \ell, e_j \rangle e_j^* \,,\, e_i \Big\rangle \\ &= \sum_j \langle \ell, e_j \rangle \cdot \langle e_j^*, e_i \rangle \;=\; \sum_j \langle \ell, e_j \rangle \cdot \delta_{ij} \;=\; \langle \ell, e_i \rangle \end{aligned}$$

for $1 \leq i \leq n$ as claimed. $\square$

The formula developed in this proof is often useful in computing dual bases.

2.2. Corollary. *If V is finite dimensional, $\mathfrak{X} = \{e_i\}$ a basis in V, and $\mathfrak{X}^* = \{e_i^*\}$ is the dual basis in V^*, then any $\ell \in V^*$ has*

$$\ell = \sum_{i=1}^{n} \langle \ell, e_i \rangle \cdot e_i^*$$

as its expansion in the $\mathfrak{X}^*$ basis.

2.3. Exercise. If $v_1 \neq v_2$ in a finite dimensional vector space V, prove that there is an $\ell \in V^*$ such that $\langle \ell, v_1 \rangle \neq \langle \ell, v_2 \rangle$. (Thus there are enough functionals in the dual V^* to distinguish vectors in V.)
Hint: It suffices to show $v_0 \neq 0 \Rightarrow \langle \ell, v_0 \rangle \neq 0$ for some $\ell \in V^*$. (Why?) Think about bases in V that involve v_0, and their duals.

Note: This result is actually true for all infinite dimensional spaces, but the proof is harder and requires "transcendental methods" involving the *Axiom of Choice*. These methods also show that every infinite dimensional space has a basis $\mathfrak{X}$ – an (infinite) set of independent vectors such that every $v \in V$ can be written as a *finite* $\mathbb{K}$-linear combination of vectors from $\mathfrak{X}$. As an example, the basic powers $\mathfrak{X} = \{1\!\!1, x, x^2, \ldots\}$ are a basis for $\mathbb{K}[x]$ in this sense. A more challenging problem is to produce a basis for $V = \mathbb{R}$ when $\mathbb{R}$ is regarded as a vector space over the field of rationals $\mathbb{Q}$. Any such *Hamel basis* for $\mathbb{R}$ is necessarily uncountable. $\square$

2.4. Example. Consider the basis $\mathbf{u}_1 = (1, 0, 1)$, $\mathbf{u}_2 = (1, -1, 0)$, $\mathbf{u}_3 = (2, 0, -1)$ in $\mathbb{R}^3$. We shall determine the dual basis vectors $\mathbf{u}_i^*$ by computing their action as functionals on an arbitrary vector $v = (x_1, x_2, x_3)$ in $\mathbb{R}^3$.

Solution: Note that $(x_1, x_2, x_3) = \sum_{k=1}^{3} x_k \mathbf{e}_k$ where $\{\mathbf{e}_k\}$ is the standard basis in $\mathbb{R}^3$. The basis $\{\mathbf{e}_k^*\}$ dual to the standard basis $\{\mathbf{e}_k\}$ has the following action:

$$\langle \mathbf{e}_k^*, (x_1, x_2, x_3) \rangle = \Big\langle \mathbf{e}_k^*, \sum_{i=1}^{3} x_i \mathbf{e}_i \Big\rangle = x_k$$

because $\mathbf{e}_k^*$ reads the k^{th} coefficient in $v = \sum_i x_i \mathbf{e}_i$. For a different basis such as $\mathfrak{Y} = \{\mathbf{u}_i\}$, the dual vector $\mathbf{u}_k^*$ reads the k^{th} coefficient c_k when we expand a typical vector $v \in \mathbb{R}^3$ as $v = \sum_{j=1}^{3} c_j \mathbf{u}_j$, so our task reduces to writing $v = (x_1, x_2, x_3) = \sum_{j=1}^{3} x_i \mathbf{e}_j$ in terms of the new basis $\{\mathbf{u}_k\}$.

In matrix form, we have:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \sum_i c_i \mathbf{u}_i = c_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}$$

so we must solve for C in the matrix equation

$$AC = X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \qquad \text{where} \qquad A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

Row operations on the augmented matrix for this system yield:

$$\begin{aligned} [A : X] \quad &= \left(\begin{array}{ccc|c} 1 & 1 & 2 & x_1 \\ 0 & -1 & 0 & x_2 \\ 1 & 0 & -1 & x_3 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 1 & 2 & x_1 \\ 0 & 1 & 0 & -x_2 \\ 0 & -1 & -3 & x_3 - x_1 \end{array}\right) \\ &\to \left(\begin{array}{ccc|c} \boxed{1} & 1 & 2 & x_1 \\ 0 & \boxed{1} & 0 & -x_2 \\ 0 & 0 & \boxed{1} & \frac{1}{3}(x_1 + x_2 - x_3) \end{array}\right) \end{aligned}$$

There are no free variables; backsolving yields the unique solution

$$\begin{aligned} c_1 &= x_1 - c_2 - 2c_3 \; = \; \tfrac{1}{3}x_1 + \tfrac{1}{3}x_2 + \tfrac{2}{3}x_3 \\ c_2 &= -x_2 \\ c_3 &= \tfrac{1}{3}(x_1 + x_2 - x_3) \end{aligned}$$

Thus,
$$v = \left(\tfrac{1}{3}x_1 + \tfrac{1}{3}x_2 + \tfrac{2}{3}x_3\right)\mathbf{u}_1 - x_2\mathbf{u}_2 + \left(\tfrac{1}{3}x_1 + \tfrac{1}{3}x_2 - \tfrac{1}{3}x_3\right)\mathbf{u}_3$$
Now read off the coefficients when $v = (x_1, x_2, x_3)$. Since $\langle \mathbf{u}_i^*, \mathbf{u}_j \rangle = \delta_{ij}$ we get

$$\begin{aligned} \langle\, \mathbf{u}_i^*, v \,\rangle &= \langle\, \mathbf{u}_i^*, (x_1, x_2, x_3)\,\rangle = \Big\langle \mathbf{u}_i^*, \sum_i x_i\mathbf{e}_i \Big\rangle \\ &= \Big\langle \mathbf{u}_i^*, \sum_j c_j\mathbf{u}_j \Big\rangle = c_i = \begin{cases} \frac{1}{3}x_1 + \frac{1}{3}x_2 + \frac{2}{3}x_3 & i=1 \\ \qquad\qquad\quad -x_2 & i=2 \\ \frac{1}{3}x_1 + \frac{1}{3}x_2 - \frac{1}{3}x_3 & i=3 \end{cases} \end{aligned}$$

Since $\langle\, \mathbf{e}_k^*, (x_1, x_2, x_3)\,\rangle = x_k$ we can also rewrite this in the form

$$\begin{aligned} \mathbf{u}_1^* &= \tfrac{1}{3}\mathbf{e}_1^* + \tfrac{1}{3}\mathbf{e}_2^* + \tfrac{2}{3}\mathbf{e}_3 \\ \mathbf{u}_2^* &= -\mathbf{e}_2^* \\ \mathbf{u}_3^* &= \tfrac{1}{3}\mathbf{e}_1^* + \tfrac{1}{3}\mathbf{e}_2^* - \tfrac{1}{3}\mathbf{e}_3^* \end{aligned}$$

by Corollary 2.2. $\square$

III.3. The Transpose Operation.

There is a natural connection between linear operators $T : V \to W$ and operators in the opposite direction, from $W^* \to V^*$.

3.1. Theorem. *The* **transpose** $T^{\mathrm{t}} : W^* \to V^*$ *of a linear operator* $T : V \to W$ *between finite dimensional vector spaces is a linear operator that is uniquely determined in a coordinate-free manner by requiring that*

$$\langle\, T^t(\ell), v\rangle = \langle \ell, T(v)\rangle \qquad \textit{for all } \ell \in W^*, v \in V \tag{26}$$

Proof: The right side of (26) defines a map $\phi_\ell : V \to \mathbb{K}$ such that $\phi_\ell(v) = \langle \ell, T(v)\rangle$. Observe that ϕ_ℓ is a linear functional on V (easily verified), so each $\ell \in W^*$ determines a well defined element of V^*. Now let $T^{\mathrm{t}} : W^* \to V^*$ be the map $T^{\mathrm{t}}(\ell) = \phi_\ell$. The property (26) holds by definition, but we must prove T^{t} is linear (and uniquely determined by the property (26)).

Uniqueness is easy: if $S : W^* \to V^*$ is another operator such that

$$\langle S(\ell), v\rangle = \langle \ell, T(v)\rangle = \langle T^{\mathrm{t}}\ell, v\rangle \qquad \text{for all } \ell \in W^* \text{ and } v \in V,$$

these identities imply $S(\ell) = T^{\mathrm{t}}(\ell)$ for all ℓ, which means $S = T^{\mathrm{t}}$ as maps on W^*.

The easiest proof that T^{t} is linear uses the scalar identities (26) and the following general observation.

3.2. Exercise. If V, W are finite dimensional vector spaces, explain why the following statements regarding two linear operators $A, B : V \to W$ are equivalent.

1. $A = B$ as operators.
2. $Av = Bv$ for all $v \in V$.
3. $\langle \ell, Av\rangle = \langle \ell, Bv\rangle$ for all $v \in V, \ell \in W^*$.

Hint: Use Exercise 2.3 to prove (3.) $\Rightarrow$ (2.); implications (2.) $\Rightarrow$ (1.) $\Rightarrow$ (3.) are trivial.

To prove $T^{\mathrm{t}}(\ell_1 + \ell_2) = T^{\mathrm{t}}(\ell_1) + T^{\mathrm{t}}(\ell_2)$ just bracket these with an arbitrary $v \in V$ and compute:

$$\begin{aligned} \langle T^{\mathrm{t}}(l_1 + l_2), v\rangle &= \langle l_1 + l_2, T(v)\rangle \\ &= \langle l_1, T(v)\rangle + \langle l_2, T(v)\rangle \qquad \text{(definition of (+) in } W^*) \\ &= \langle T^{\mathrm{t}}(l_1) + T^{\mathrm{t}}(l_2), v\rangle \qquad \text{(definition of (+) in } V^*) \end{aligned}$$

for all $v \in V$. The other identity we need, $T^{\rm t}(\lambda \cdot \ell) = \lambda \cdot T^{\rm t}(\ell)$, is proved similarly. $\square$

Thus $T^{\rm t} : W^* \to V^*$ is a well-defined linear operator that acts in the opposite direction from $T : V \to W$.

Basic properties of the correspondence $T \to T^{\rm t}$ are left as exercises. The proofs are easy using the scalar identities (26).

3.3. Exercise. Verify that

1. The transpose $0^{\rm t}$ of the zero operator $0(v) \equiv 0_W$ from $V \to W$ is the zero operator from $W^* \to V^*$, so $0^{\rm t}(\ell) = 0_{V^*}$ for all $\ell \in W^*$.
2. When $V = W$ the transpose of the identity map $\mathrm{id}_V : V \to V$, with $\mathrm{id}_V(v) \equiv v$, is the identity map $\mathrm{id}_{V^*} : V^* \to V^*$ – in short, $\big(\mathrm{id}_V\big)^{\rm t} = \mathrm{id}_{V^*}$.
3. $(\lambda_1 T_1 + \lambda_2 T_2)^{\rm t} = \lambda_1 T_1^{\rm t} + \lambda_2 T_2^{\rm t}$, for any $\lambda_1, \lambda_2 \in \mathbb{K}$ and $T_1, T_2 : V \to W$.

3.4. Exercise. If $U \xrightarrow{T} V \xrightarrow{S} W$ are linear maps between finite dimensional vector spaces, prove that

$$\big(S \circ T\big)^{\rm t} = T^{\rm t} \circ S^{\rm t}$$

Note the reversal of order when we compute the transpose of a product.

3.5. Exercise. If V, W are finite dimensional and $T : V \to W$ is an invertible linear operator (a bijection), prove that $T^{\rm t} : W^* \to V^*$ is invertible too, and $\big(T^{-1}\big)^{\rm t} = \big(T^{\rm t}\big)^{-1}$ as maps from $V^* \to W^*$.

Now for some computational issues

3.6. Theorem. *Let $T : V \to W$ be a linear operator between finite dimensional spaces, let $\mathfrak{X} = \{v_1, ..., v_m\}$, $\mathfrak{Y} = \{w_1, ..., w_n\}$ be bases in V, W and let $\mathfrak{X}^* = \{v_i^*\}$, $\mathfrak{Y}^* = \{w_j^*\}$ be the dual bases in V^*, W^*. We have defined the transpose $A^{\rm t}$ of an $n \times m$ matrix to be the $m \times n$ matrix such that $(A^{\rm t})_{ij} = A_{ji}$. Then "$[T^{\rm t}] = [T]^{\rm t}$" in the sense that*

$$[T^{\rm t}]_{\mathfrak{X}^*\mathfrak{Y}^*} = \big([T]_{\mathfrak{Y}\mathfrak{X}}\big)^{\rm t}$$

Important Note: This only works for the dual bases $\mathfrak{X}^*$, $\mathfrak{Y}^*$ in V^*, W^*. If $\mathfrak{A}, \mathfrak{B}$ are arbitrary bases in V^*, W^* unrelated to the dual bases there is no reason to expect that

$$[T^{\rm t}]_{\mathfrak{B}\mathfrak{A}} = \text{ the transpose of the matrix } [T]_{\mathfrak{Y}\mathfrak{X}} \qquad \square$$

Proof: To determine $[T^{\rm t}]$ we must calculate the coefficients $[T^{\rm t}]_{ji}$ in the system of vector equations

$$T^{\rm t}(w_i^*) = \sum_{j=1}^{m} [T^{\rm t}]_{ji} v_j^* \qquad 1 \le i \le n$$

These are easily found by applying each of these identities to a basis vector v_k in V:

$$\langle T^{\rm t}(w_i^*), v_k\rangle = \sum_{j=1}^{m} [T^{\rm t}]_{ji} \cdot \langle v_j^*, v_k\rangle = \sum_{j=1}^{m} [T^{\rm t}]_{ji} \delta_{jk} = [T^{\rm t}]_{ki} \tag{27}$$

for any $1 \le i \le n$ and $1 \le k \le m$. Thus

$$[T^{\rm t}]_{ki} = \langle T^{\rm t}(w_i^*), v_k\rangle$$

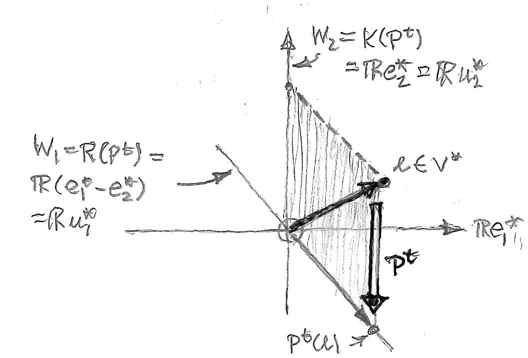

Figure 3.3. The decomposition $V = W_1 \oplus W_2$ determines the projection $P : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ in Example 3.7 that maps V onto $W_1 = \mathbb{R}\mathbf{u}_1$ along $W_2 = \mathbb{R}\mathbf{u}_2$. We show that its transpose P^{t} projects V^* onto its range $R(P^{\mathrm{t}}) = \mathbb{R}\mathbf{u}_1^* = \mathbb{R}(\mathbf{e}_1^* - \mathbf{e}_2^*)$, along $K(P^{\mathrm{t}}) = \mathbb{R}\mathbf{u}_2^* = \mathbb{R}\mathbf{e}_2^*$. Horizontal axis in this picture is $\mathbb{R}\mathbf{e}_1^*$ and vertical axis is $\mathbb{R}\mathbf{e}_2^*$; a functional is then represented as $\ell = (\dot{x}_1\mathbf{e}_1^* + \dot{x}_2\mathbf{e}_2^*)$ with respect to the basis $\mathfrak{X}^* = \{\mathbf{e}_1^*, \mathbf{e}_2^*\}$ dual to the standard basis $\mathfrak{X} = \{\mathbf{e}_1, \mathbf{e}_2\}$.

By definition of T^{t} and the matrix $[T]_{\mathfrak{Y}\mathfrak{X}}$, we can also write (27) as

$$\begin{aligned} \langle T^{\mathrm{t}}(w_i^*), v_k\rangle &= \langle w_i^*, T(v_k)\rangle = \Big\langle w_i^*, \sum_{j=1}^{n} [T]_{jk} w_j \Big\rangle \\ &= \sum_{j=1}^{j} [T]_{jk}\langle w_i^*, w_j\rangle = \sum_{j=1}^{n} [T]_{jk}\delta_{ij} = [T]_{ik} \end{aligned}$$

for any $1 \leq i \leq n$, $1 \leq k \leq m$.

Upon comparison with previous result we conclude that $[T^{\mathrm{t}}]_{ki} = [T]_{ik} = \big([T]^{\mathrm{t}}\big)_{ki}$. Thus $[T^{\mathrm{t}}]_{\mathfrak{X}^*\mathfrak{Y}^*}$ is the transpose of $[T]_{\mathfrak{Y}\mathfrak{X}}$. $\square$

3.7. Exercise (Computing Matrix Entries). If $T : V \rightarrow W$ and bases $\mathfrak{X} = \{e_i\}$, $\mathfrak{Y} = \{f_i\}$ are given in V, W let $\mathfrak{X}^*$, $\mathfrak{Y}^*$ be the dual bases. Prove that

$$[T]_{\mathfrak{Y}\mathfrak{X}} = [t_{ij}] \text{ has entries } t_{ij} = \langle f_j^*, T(e_i)\rangle$$

The transpose of a projection $P : V \rightarrow V$ is a projection $P^{\mathrm{t}} : V^* \rightarrow V^*$ because $P^{\mathrm{t}} \circ P^{\mathrm{t}} = (P \circ P)^{\mathrm{t}} = P^{\mathrm{t}}$, so P^{t} maps V^* onto the range $R(P^{\mathrm{t}})$ along the nullspace $K(P^{\mathrm{t}}) = \ker(P^{\mathrm{t}})$ in the direct sum $V^* = R(P^{\mathrm{t}}) \oplus K(P^{\mathrm{t}})$. The following example shows how to calculate these geometric objects in terms of dual bases.

3.8. Example. Let $V = \mathbb{R}^2$ with basis $\mathfrak{Y} = \{\mathbf{u}_1, \mathbf{u}_2\}$ where $\mathbf{u}_1 = (1, 0)$, $\mathbf{u}_2 = (1, 1)$, and let $P =$ projection onto $W_1 = \mathbb{R}\mathbf{u}_1$ along $W_2 = \mathbb{R}\mathbf{u}_2$. The standard basis $\mathfrak{X} = \{\mathbf{e}_1, \mathbf{e}_2\}$ or the basis $\mathfrak{Y} = \{\mathbf{u}_1, \mathbf{u}_2\}$ can be used to describe P. The description with respect to $\mathfrak{X}$ has already been worked out in Example 3.6 of Chapter II.

1. Compute the dual bases $\mathfrak{X}^*$, $\mathfrak{Y}^*$ as functions $\ell : \mathbb{R}^2 \rightarrow \mathbb{R}$ and find the matrix descriptions of P^{t}:
$$[P^{\mathrm{t}}]_{\mathfrak{X}^*\mathfrak{X}^*} \qquad \text{and} \qquad [P^{\mathrm{t}}]_{\mathfrak{Y}^*\mathfrak{Y}^*}$$
2. Compute the kernel $K(P^{\mathrm{t}})$ and the range $R(P^{\mathrm{t}})$ in terms of the basis $\mathfrak{X}^*$ dual to the standard basis $\mathfrak{X}$.

3. Repeat (2.) for the basis $\mathfrak{Y}^*$.

Solution: First observe that

$$\begin{cases} \mathbf{u}_1 &= \mathbf{e}_1 \\ \mathbf{u}_2 &= \mathbf{e}_1 + \mathbf{e}_2 \end{cases} \quad \Rightarrow \quad \begin{cases} \mathbf{u}_1 &= \mathbf{e}_1 \\ \mathbf{e}_2 &= \mathbf{u}_2 - \mathbf{u}_1 \end{cases}$$

By definition, $\mathfrak{Y} = \{\mathbf{u}_1, \mathbf{u}_2\}$ is a diagonalizing basis for P, with

$$\begin{cases} P(\mathbf{u}_1) &= \mathbf{u}_1 \\ P(\mathbf{u}_2) &= \mathbf{0} \end{cases} \qquad \text{which } \Rightarrow \quad [P]_{\mathfrak{Y}\mathfrak{Y}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

We also have

$$\begin{cases} P(\mathbf{e}_1) &= P(\mathbf{u}_1) = \mathbf{u}_1 = \mathbf{e}_1 \\ P(\mathbf{e}_2) &= P(\mathbf{u}_2 - \mathbf{u}_1) = -\mathbf{u}_1 = -\mathbf{e}_1 \end{cases} \qquad \text{which } \Rightarrow \quad [P]_{\mathfrak{X}\mathfrak{X}} = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$$

The dual basis vectors are computed as functions $\mathbb{R}^2 \to \mathbb{R}$ by observing that

$$\begin{aligned} \mathbf{u}_1^*(v_1, v_2) &= \mathbf{u}_1^*(v_1\mathbf{e}_1 + v_2\mathbf{e}_2) \\ &= \mathbf{u}_1^*\big((v_1 - v_2)\mathbf{u}_1 + v_2\mathbf{u}_2\big) \qquad \text{which } \Rightarrow \mathbf{u}_1^* = \mathbf{e}_1^* - \mathbf{e}_2^* \\ &= v_1 - v_2 = (\mathbf{e}_1^* - \mathbf{e}_2^*)(v_1, v_2) \end{aligned}$$

and

$$\begin{aligned} \mathbf{u}_2^*(v_1, v_2) &= \mathbf{u}_2^*(v_1\mathbf{e}_1 + v_2\mathbf{e}_2) \\ &= \mathbf{u}_2^*\big((v_1 - v_2)\mathbf{u}_1 + v_2\mathbf{u}_2\big) \qquad \text{which } \Rightarrow \mathbf{u}_2^* = \mathbf{e}_2^* \\ &= v_2 = \mathbf{e}_2^*(v_1, v_2) \end{aligned}$$

No further calculations are needed to finish (1.); just apply Theorem 3.6 to get

$$[P^{\mathrm{t}}]_{\mathfrak{X}^*\mathfrak{X}^*} = \big([P]_{\mathfrak{X}\mathfrak{X}}\big)^{\mathrm{t}} = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}$$

Applying the same idea we see that

$$[P^{\mathrm{t}}]_{\mathfrak{Y}^*\mathfrak{Y}^*} = \big([P]_{\mathfrak{Y}\mathfrak{Y}}\big)^{\mathrm{t}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}^{\mathrm{t}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad .$$

That resolves Question 1.

For (2.), a functional $\ell = \dot{x}_1\mathbf{e}_1^* + \dot{x}_2\mathbf{e}_2^*$ ($\dot{x}_i \in \mathbb{R}$) is in $K(P^{\mathrm{t}}) \Leftrightarrow$

$$[P^{\mathrm{t}}]_{\mathfrak{X}^*\mathfrak{X}^*} \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} \dot{x}_1 \\ -\dot{x}_1 \end{pmatrix} \text{ is equal to } \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

That happens $\Leftrightarrow \dot{x}_1 = 0$, so $K(P^{\mathrm{t}}) = \mathbb{R}\mathbf{e}_2^*$ with respect to the $\mathfrak{X}^*$ basis. Since we know $\mathbf{e}_2^* = \mathbf{u}_2^*$ we get $K(P^{\mathrm{t}}) = \mathbb{R}\mathbf{u}_2^*$ with respect to the $\mathfrak{Y}^*$ basis.

As for $R(P^{\mathrm{t}})$, if $\ell = b_1\mathbf{e}_1^* + b_2\mathbf{e}_2^*$ in the $\mathfrak{X}^*$-basis and we let $B = \mathrm{col}(b_1, b_2)$, we must solve the matrix equation $A\dot{X} = B$, where $A = [P^{\mathrm{t}}]_{\mathfrak{X}^*\mathfrak{X}^*}$. Row operations on $[A : B]$ yield

$$\left(\begin{array}{cc|c} 1 & 0 & b_1 \\ -1 & 0 & b_2 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & 0 & b_1 \\ 0 & 0 & b_2 + b_1 \end{array}\right)$$

so $B \in R(P^{\mathrm{t}}) \Leftrightarrow b_1 + b_2 = 0 \Leftrightarrow \ell \in \mathbb{R}\cdot(\mathbf{e}_1^* - \mathbf{e}_2^*)$. Thus $R(P^{\mathrm{t}}) = \mathbb{R}\cdot(\mathbf{e}_1^* - \mathbf{e}_2^*)$ in the $\mathfrak{X}^*$ basis, while in the $\mathfrak{Y}^*$ basis this becomes

$$R(P^{\mathrm{t}}) = \mathbb{R}\big(\mathbf{e}_1^* - \mathbf{e}_2^*\big) = \mathbb{R}\big((\mathbf{u}_2^* + \mathbf{u}_1^*) - \mathbf{u}_2^*\big) = \mathbb{R}\mathbf{u}_1^* \quad \square$$

The projection P^{t}, and the corresponding decomposition $V^* = R(P^{\mathrm{t}}) \oplus K(P^{\mathrm{t}})$, both have coordinate-independent geometric meaning. But the components of the direct sum have different descriptions according to which dual basis we use to describe vectors in V^*:

$$V^* = R(P^{\rm t}) \oplus K(P^{\rm t}) = \begin{cases} \mathbb{R}(\mathbf{u}_1^*) \oplus \mathbb{R}(\mathbf{u}_2^*) & \text{for the } \mathfrak{Y}^* \text{ basis} \\ \mathbb{R}(\mathbf{e}_1^* - \mathbf{e}_2^*) \oplus \mathbb{R}(\mathbf{e}_2^*) & \text{for the } \mathfrak{X}^* \text{ basis} \end{cases} \qquad \Box$$

3.9. Exercise. Round out the previous discussion by verifying that

1. For the standard basis $\mathfrak{X}$ we have $R(P) = \mathbb{R}\mathbf{e}_1$ and $K(P) = \mathbb{R}\cdot(\mathbf{e}_1 + \mathbf{e}_2)$.
2. For the $\mathfrak{Y}$ basis we have $R(P) = \mathbb{R}\mathbf{u}_1$ and $K(P) = \mathbb{R}(\mathbf{u}_2)$.

Reflexivity of Finite Dimensional Spaces. If V is finite dimensional there is a natural "bracketing map"

$$\phi : V^* \times V \to \mathbb{K} \qquad \text{given by} \qquad \phi : (\ell, v) \mapsto \langle \ell, v\rangle$$

The expression $\langle \ell, v\rangle$ is linear in each variable when the other is held fixed. If ℓ is fixed we get a linear functional $v \mapsto \ell(v)$ on V, but if we fix v the map $\ell \mapsto \langle \ell, v\rangle$ is a linear map from $V^* \to \mathbb{K}$, and hence is an element $j(v) \in V^{**} = (V^*)^*$, the "double dual" of V.

3.10. Lemma. *If* $\dim(V) < \infty$ *the map* $j : V \to V^{**}$ *is linear and a bijection. It is a "natural" isomorphism* (*defined without reference to any coordinate system*) *that allows us to identify* V^{**} *with* V*.*

Proof: For any $\ell \in V^*$ we have

$$\langle j(v_1 + v_2), \ell\rangle = \langle \ell, v_1 + v_2\rangle = \langle \ell, v_1\rangle + \langle \ell, v_2\rangle = \langle j(v_1), \ell\,\rangle + \langle j(v_2), \ell\,\rangle$$

and similarly

$$\langle j(\lambda\cdot v), \ell\rangle = \langle \ell, \lambda v\rangle = \lambda\cdot\langle \ell, v\rangle = \langle \lambda\cdot j(v), \ell\,\rangle$$

Since these relations are true for all $\ell \in V^*$ we see that $j(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 j(v_1) + \lambda_2 j(v_2)$ in V^{**} and $j : V \to V^{**}$ is linear.

Finite dimensionality of V insures that $\dim(V^{**}) = \dim(V^*) = \dim(V)$, so j is a bijection $\Leftrightarrow$ j is onto $\Leftrightarrow$ j is one-to-one $\Leftrightarrow$ $\ker(j) = (0)$. But $j(v) = 0$ if and only if $0 = \langle j(v), \ell\,\rangle = \langle \ell, v\rangle$ for every $\ell \in V^*$. This forces $v = 0$ (and hence $\ker(j) = (0)$) because if $v \neq 0$ there is a functional $\ell \in V^*$ such that $\langle \ell, v\rangle \neq 0$. [In fact, we can extend $\{v\}$ to a basis $\{v, v_2, ..., v_n\}$ of V. Then, if we form the dual basis $\{v^*, v_2^*, ..., v_n^*\}$ in V^* we have $\langle v^*, v\rangle = 1$.] $\Box$

There is, on the other hand, no *natural* (basis-independent) isomorphism from V to V^*. The spaces V and V^* are isomorphic because they have equal dimension, so there are many un-natural bijective linear maps between them. (We can create such a map given any basis $\{e_i\} \subseteq V$ and any basis $\{f_j\} \subseteq V^*$ by sending $e_i \to f_i$.)

If we identify $V = V^{**}$ via the natural map j, then the dual basis $(\mathfrak{X}^*)^*$ gets identitfied with the original basis $\mathfrak{X}$ in V. [Details: In fact, the vector e_i^{**} in the dual basis to $\mathfrak{X}^*$ coincides with the image vector $j(e_i)$ because

$$\langle j(e_i), e_j^*\rangle = \langle e_j^*, e_i\rangle = \delta_{ij}\ ,$$

which is the defining property of the vectors $(e_i^*)^*$ in $\mathfrak{X}^{**}$. Hence $j(e_i) = e_i^{**}$.] By time-honored abuse of notation mathematicians often write "$\mathfrak{X}^{**} = \mathfrak{X}$" even though this is not strictly true.

Furthermore when we identify $V^{**} \cong V$, the "double transpose" $T^{\rm tt} = (T^{\rm t})^{\rm t}$ mapping $V^{**} \to V^{**}$ becomes the original operator T, allowing us to write

$$T^{\rm tt} = T \qquad \text{(again, by abuse of notation)}$$

The precise connection between T and T^{tt} is shown in the following commutative diagram

$$\begin{array}{ccc} V^{**} & \xrightarrow{(T^{\mathrm{t}})^{\mathrm{t}}} & V^{**} \\ j^{-1}\downarrow\uparrow j & & \uparrow j \\ V & \xrightarrow{T} & V \end{array} \qquad \big(\text{Diagram commutes:} \quad T^{\mathrm{tt}} = j \circ T \circ j^{-1}\big)$$

3.11. Exercise. If $|V| = \dim V < \infty$, $\mathfrak{X} = \{e_1, ..., e_n\}$ is a basis, and $T : V \to V$ a linear operator,

1. Fill in the details needed to show that the diagram above commutes,
$$(T^{\mathrm{t}})^{\mathrm{t}} \circ j = j \circ T$$

2. Prove the following useful fact relating matrix realizations of T and T^{tt}:
$$[T^{\mathrm{tt}}]_{\mathfrak{X}^{**}\mathfrak{X}^{**}} = [T]_{\mathfrak{X}\mathfrak{X}}$$
for the bases $\mathfrak{X}$ and $\mathfrak{X}^{**} = j(\mathfrak{X})$.

For infinite dimensional spaces there is still a natural linear embedding $j : V \to V^{**}$. Although j is again one-to-one, it is not necessarily onto and there is a chain of distinct dual spaces V, V^* , V^{**}, V^{***}, ... When $\dim(V) < \infty$, this process terminates with $V^{**} \cong V$. For this reason finite dimensional vector spaces are said to be "*reflexive.*" (Some infinite dimensional space are reflexive too, but not many.)

Annhilators. Additional structure must imposed on a vector space in order to speak of "lengths" or "orthogonality" of vectors, or the "orthogonal complement" $W^{\perp}$ of some subspace. When $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, this is most often accomplished by imposing an "inner product" $B : V \times V \to \mathbb{K}$ on the space. However, in the absence of such extra structure there is still a natural notion of a "complementary subspace" to any subspace $W \subseteq V$; but this complement

$$W^{\circ} = \{\ell \in V^* : \langle \ell, w\rangle = 0 \text{ for all } w \in W\} \qquad (\text{the } \textbf{annihilator of } W)$$

lives in V^* rather than V. It is easily seen that W° is a vector subspace in V^*. Obviously $(0)^{\circ} = V^*$ and $V^{\circ} = (0)$ in V^*, and when W is a proper subspace in V the annihilator W° is a proper subspace of V^*, with $(0) \subsetneqq W^{\circ} \subsetneqq V^*$.

3.12. Lemma. *Let V be finite dimensional and $W \subsetneqq V$ a subspace. If $v_0 \in V$ lies outside of W there is a functional $\ell \in V^*$ such that $\langle \ell, W\rangle = 0$ so $\ell \in W^{\circ}$ but $\langle \ell, v_0\rangle \neq 0$.*

We leave the proof as an exercise. (The idea is simply to take a basis $\{e_1, ..., e_r\}$ for W (with $W \neq V$ and $r < n = \dim(V)$. Given a vector $v_0 \notin W$, adjoin additional vectors $e_{r+1} = v_0$, e_{r+2}, ..., e_n to make a basis $\mathfrak{X}$ for V. The dual basis $\mathfrak{X}^*$ provides the answer.)

We list the basic properties of annihilators as a series of exercises, some of which are major theorems (hints provided). In proving any of these results you may use any prior exercise. In all cases we assume $\dim(V) < \infty$.

3.13. Exercise. Let W be a subspace and $\mathfrak{X} = \{e_1, \ldots, e_r, \ldots, e_n\}$ a basis for V such that $\{e_1, \ldots, e_r\}$ is a basis in W. If $\mathfrak{X}^* = \{e_i^*\}$ is the dual basis, prove that $\{e_{r+1}^*, \ldots, e_n^*\}$ is a basis for the annihilator $W^{\circ} \subseteq V^*$.

3.14. Exercise. (Dimension Theorem for Annihilators). If W is a subspace in a finite dimensional vector space V, prove that

$$\dim_{\mathbb{K}}(W) + \dim_{\mathbb{K}}(W^{\circ}) = \dim_{\mathbb{K}}(V) \quad , \tag{28}$$

or in abbreviated form, $|W| + |W^{\circ}| = |V|$.

3.15. Lemma. If $T : V \to W$ is a linear operator,

1. Prove that
$$K(T^{\mathrm{t}}) = R(T)^{\circ} \qquad \big(\text{annihilator of the range } R(T)\big)$$
2. Is it also true that $R(T^{\mathrm{t}}) = K(T)^{\circ}$? If not, what goes wrong?

3.16. Exercise. If $T : V \to V$ is linear and W a subspace of V, prove that W is T-invariant if and only if its annihilator W° is invariant under the transpose T^{t}.

3.17. Exercise. If $T : V \to W$ is linear, prove that $\mathrm{rank}(T^{\mathrm{t}}) = \mathrm{rank}(T)$.

Recall that the rank of any linear operator $T : V \to W$ is the dimension $|R(T)| = \dim(R(T))$ of its range, and if $A \in \mathrm{M}(n \times m, \mathbb{K})$ we defined $L_A : \mathbb{K}^m \to \mathbb{K}^n$ via $L_A(v) = A \cdot v$, for $v \in \mathbb{K}^m$. Furthermore, recall that the "column rank" of a matrix is the dimension of its column space: $\mathrm{colrank}(A) = \dim\big(\mathrm{Col}(A)\big)$, and similarly $\mathrm{rowrank}(A) = \dim\big(\mathrm{Row}(A)\big)$. It is important to know that these numbers, computed in entirely different ways, are always equal – i.e. "*row rank = column rank*" for any matrix, regardless of its shape. The following exercises address this issue.

3.18. Exercise. If $T : V \to W$ is a linear map between finite dimensional spaces, $\mathfrak{X} = \{e_i\}$, $\mathfrak{Y} = \{f_j\}$ are bases, and $A = [T]_{\mathfrak{Y}\mathfrak{X}}$, prove that

1. The range $R(L_A)$ is equal to column space $\mathrm{Col}(A)$, hence
$$\mathrm{rank}(L_A) = \dim(R(L_A)) = \dim\big(\mathrm{Col}(A)\big) = \mathrm{colrank}(A)$$
for any $n \times m$ matrix.
2. If $A = [T]_{\mathfrak{Y},\mathfrak{X}}$ then $\mathrm{rank}(T) = \mathrm{rank}(L_A) = \mathrm{colrank}(A)$

Hint: For (2.) recall the commutative diagram Figure 2.3 of Chapter II; the vertical maps are isomorphisms and isomorphisms preserve dimensions of subspaces.

3.19. Exercise. If A is an $n \times m$ matrix,

1. Prove that $\mathrm{rank}\big(L_{A^{\mathrm{t}}}\big) = \mathrm{rank}\big((L_A)^{\mathrm{t}}\big)$.
2. This would follow if it were true that "$(L_A)^{\mathrm{t}} = L_{A^{\mathrm{t}}}$." Explain why this statement does not make sense.

Hint: Keep in mind the setting for this (and the next) Exercise. If $V = \mathbb{K}^m, W = \mathbb{K}^n$, and A is $n \times m$ we get a map $L_A : \mathbb{K}^m \to \mathbb{K}^n$. The transpose A^{t} is $m \times n$ and determines a linear map in the opposite direction:

$$V \xrightarrow{L_A} W \qquad V \xleftarrow{L_{A^{\mathrm{t}}}} W \qquad V^* \xleftarrow{(L_A)^{\mathrm{t}}} W^*$$

The transpose $\big(L_A\big)^{\mathrm{t}}$ maps $W^* \to V^*$.

Use the results of the previous exercises to prove the main result below.

3.20. Exercise. If A is an $n \times m$ matrix, prove that

Theorem: *For any $n \times m$ matrix,* $\mathrm{rowrank}(A) = \mathrm{colrank}(A)$