<u>INDEX CONVENTIONS</u> We need an efficient abbreviated notation to handle the complexity of the mathematical structure before us. We will use indices of a given "type" to denote all possible values of given index ranges. By index type we mean

$a, b, c, \ldots$

$i, j, k, \ldots$

$\alpha, \beta, \gamma, \ldots$

Variations might be barred or primed letters or capital letters.

For example, suppose we are looking at linear transformations between $\mathbb{R}^n$ and $\mathbb{R}^m$. We would need two different index ranges, say $i, j, k, \ldots = 1, 2, \ldots, n$ and $\alpha, \beta, \gamma, \ldots = 1, 2, \ldots, m$.

A given index letter should only occur once in a given expression, in which the expression stands for all expressions for which the index assumes its allowed values (call this a "free index"), or twice but only as a superscript subscript pair (one up, one down) which will stand for the sum over all allowed values (call this a "repeated index")

$A^i \leftrightarrow$ n expressions: $A^1, A^2, \ldots, A^n$

$A^i{}_i \leftrightarrow \sum_{i=1}^{n} A^i{}_i$ single expression with n terms

$A^{ji}{}_i \leftrightarrow \sum_{i=1}^{n} A^{1i}{}_i, \ldots, \sum_{i=1}^{n} A^{ni}{}_i$ n expressions each of which has n terms in the sum

$A_{ii} \leftrightarrow$ no sum, just an expression for each i, or if we want to refer to a specific diagonal component (entry) of a matrix, for example.

$A^i(v_i + w_i) = A^i v_i + A^i w_i$ ALLOWED EXCEPTION since meaning clear

A repeated index is a "dummy index" (like the dummy variable in a definite integral $\int_a^b f(x)\,dx = \int_a^b f(u)\,du$): $A^i{}_i = A^j{}_j$. We can change them at will.

A Vector Space and its Dual Space.

Let V be a finite dimensional real vector space, $\dim V = n$. Elements of this space are called "vectors". A basis of V $\{e_i\}_{i=1,\dots,n}$ or just $\{e_i\}$ is a linearly independent spanning set for V.

1) spanning condition: any vector $v \in V$ can be represented as a linear combination of the basis vectors:

$$v = \sum_{i=1}^{n} \underbrace{v^i}_{\text{coefficients: "components" of } v \text{ wrt } \{e_i\}} e_i = v^i e_i$$

this index labels components

this one labels basis vectors

2) linear independence:

if $v^i e_i = 0$, then $v^i = 0$ $\left[\text{If } \sum_{i=1}^{n} v^i e_i = 0, \text{ then } v^i = 0 \text{ for all } i=1,\dots,n \right]$

EX $V = \mathbb{R}^n = \{ u = (u^1, \dots, u^n) = (u^i) \mid u^i \in \mathbb{R} \}$ ~ space of n-tuples of real numbers

natural basis $e_1 = (1,0,\dots,0)$, $e_2 = (0,1,0,\dots,0)$, $\dots$, $e_n = (0,\dots,0,1)$.

Let V^* be the "dual space" of V, equal to the space of real-valued linear functions on V; elements of V^* called "covectors."

condition of linearity: $f \in V^* \rightarrow f(au + bv) = af(u) + bf(v)$ (a, b real numbers!)

"*value of linear combination = linear combination of values.*"

This easily extends to linear combinations with any number of terms; for example:

$$f(v) = f\left(\sum_{i=1}^{n} v^i e_i\right) = \sum_{i=1}^{n} v^i \underbrace{f(e_i)}_{\equiv f_i}$$

f_i: "components" of covector wrt basis $\{e_i\}$

or equivalently $f(v^i e_i) = v^i f(e_i) = v^i f_i$.

A covector is entirely determined by its values on the basis vectors.

The dual space V^* is itself an n-dimensional vector space, with linear combinations of covectors defined in the usual way one takes

A Vector Space and its Dual Space.

Let V be a finite dimensional real vector space, $\dim V = n$. Elements of this space are called "vectors". A basis of V $\{e_i\}_{i=1,\dots,n}$ or just $\{e_i\}$ is a linearly independent spanning set for V.

1) spanning condition: any vector $v \in V$ can be represented as a linear combination of the basis vectors:

$$v = \sum_{i=1}^{n} v^i e_i = v^i e_i$$

this index labels components

this one labels basis vectors

coefficients: "components" of v wrt $\{e_i\}$

2) linear independence:

if $v^i e_i = 0$, then $v^i = 0$ $\left[\text{If } \sum_{i=1}^{n} v^i e_i = 0, \text{ then } v^i = 0 \text{ for all } i = 1,\dots,n\right]$

EX $V = \mathbb{R}^n = \{ u = (u^1, \dots, u^n) = (u^i) \mid u^i \in \mathbb{R} \}$ ~ space of n-tuples of real numbers

natural basis $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, $\dots$, $e_n = (0, \dots, 0, 1)$.

Let V^* be the "dual space" of V, equal to the space of real-valued linear functions on V; elements of V^* called "covectors."

condition of linearity: $f \in V^* \longrightarrow f(au + bv) = af(u) + bf(v)$ (a, b real numbers!)

" *value of linear combination = linear combination of values.* "

This easily extends to linear combinations with any number of terms; for example:

$$f(v) = f\Big(\sum_{i=1}^{n} v^i e_i\Big) = \sum_{i=1}^{n} v^i \underbrace{f(e_i)}_{\equiv f_i}$$

f_i ← "components" of covector wrt basis $\{e_i\}$

or equivalently $f(v^i e_i) = v^i f(e_i) = v^i f_i$.

A covector is entirely determined by its values on the basis vectors.

The dual space V^* is itself an n-dimensional vector space, with linear combinations of covectors defined in the usual way one takes linear combinations of any functions, i.e., in terms of the values:

$$(af+bg)(v) \equiv af(v)+bg(v)$$

f, g covectors
v vector

Exercise: Show that this defines a linear function $af+bg$, so that the space is closed under this linear combination operation. [All the other vector space properties are inherited from the linear structure of V.]

Let us produce a basis for V^*, called the dual basis $\{\omega^i\}$ or "the basis dual to $\{e_i\}$", by finding n covectors which satisfy the following "duality relations"

$$\omega^i(e_j) = \delta^i{}_j \equiv \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}$$

"*Kronecker delta*"

If we can do this, then by linearity

$$\omega^i(v) = \omega^i(v^j e_j) = v^j \omega^i(e_j) = v^j \delta^i{}_j = v^i$$

(only term with $j=i$ contributes to sum!)

so the ith dual basis vector picks out the ith component of a vector.

Why is this a basis of V^*?

1) spanning condition:

$$\begin{aligned} f(v) &= f(v^i e_i) = v^i f(e_i) \equiv v^i f_i \\ &= v^i \delta^j{}_i f_j = v^i \omega^j(e_i) f_j = (f_j \omega^j)(v^i e_i) \\ &= (f_i \omega^i)(v) \end{aligned}$$

f and $f_i \omega^i$ have same value on every $v \in V$ so are same function:

$$f = f_i \omega^i \qquad \text{where } f_i = f(e_i)$$

this index labels basis covectors

this index labels components

"components" of f wrt basis $\{\omega^i\}$ of V^* are just the "components of f wrt the basis $\{e_i\}$" already introduced above

2) linear independence: Suppose $f_i \omega^i = 0$ = zero covector!

Then evaluating each side of this equation on e_j:

$$0 = 0(e_j) = (f_i \omega^i)(e_j) = f_i \omega^i(e_j) = f_i \delta^i{}_j = f_j,$$

so they are linearly independent.

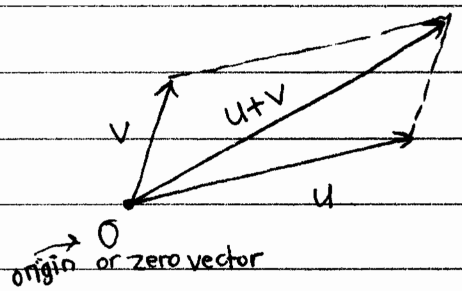
Thus V^* is also an n-dimensional vector space.

EX. The familiar cartesian coordinates on $\mathbb{R}^n$ are defined by

$$x^i((u^1,\dots,u^n)) = u^i \qquad \text{[value of ith number in n-tuple]}$$

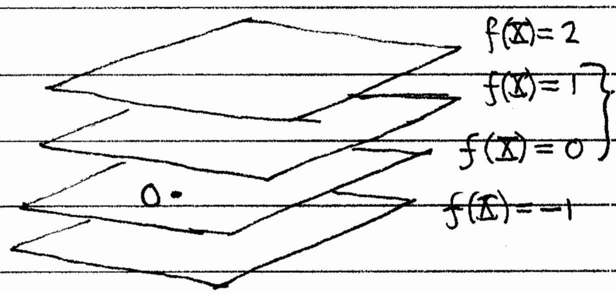
But this is exactly what the basis $\{\omega^i\}$ dual to the natural basis $\{e_i\}$ does – i.e., the set of cartesian coordinates $\{x^i\}$, interpreted as linear functions on the vector space $\mathbb{R}^n$ [why are they linear?], is the dual basis: $\omega^i = x^i$. A general linear function on $\mathbb{R}^n$ has the familiar form $f = f_i\omega^i = f_i x^i$.

Vectors and vector addition are best visualized by interpreting points in $\mathbb{R}^n$ as directed line segments. Functions can be visualized in terms of their level surfaces. For linear functions, the level surfaces



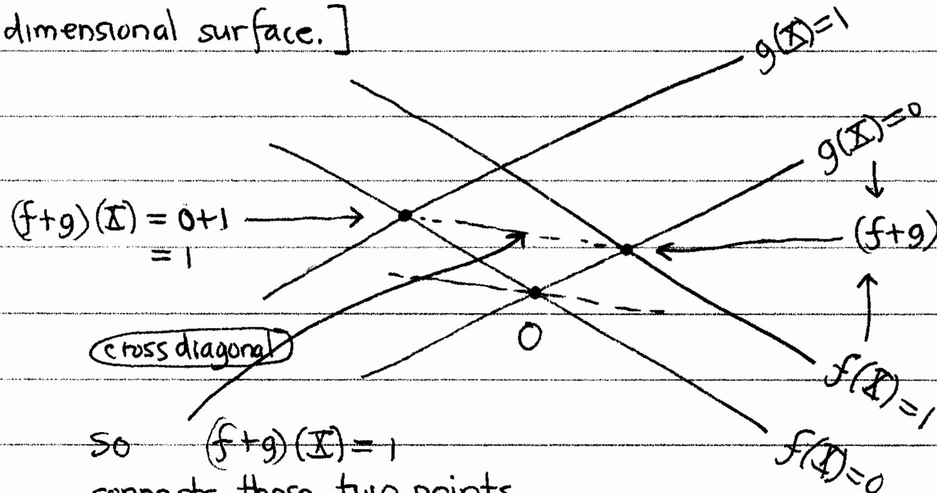
$$f_i x^i = t \qquad \text{(a parameter)}$$

are a family of parallel hyperplanes, best represented by selecting an equally spaced set of such hyperplanes, say by choosing integer values of the parameter t.



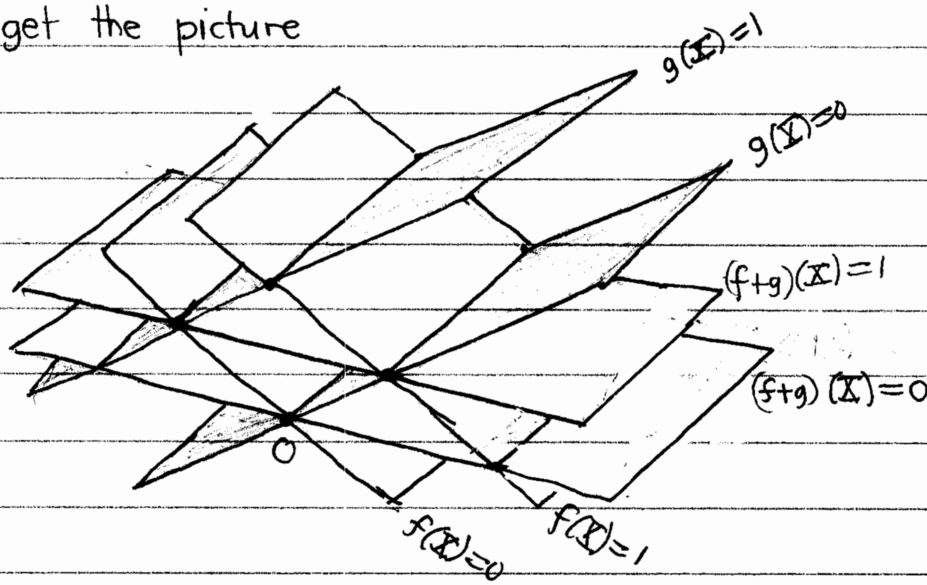
However, it is enough to graph two such level surfaces $f(X)=0$ and $f(X)=1$ to have a mental picture of the entire family. This pair of planes also enables one to have a geometric interpretation of covector addition on the vector space itself, like the parallelogram law for vectors. However, instead of the main diagonal line segment, one has the cross diagonal hyperplane for the result.

Let's look at two pairs of such hyperplanes representing f and g but "edgeon", namely in the 2-plane orthogonal to the $(n-2)$ planes of intersection of the $(n-1)$-planes which are these hyperplanes. [Recall that 2 surfaces of dimension m in general intersect in an $(m-1)$ dimensional surface.]



So $(f+g)(X)=1$
connects these two points
and $(f+g)(X)=0$ is parallel and goes through the origin.

Thus we get the picture



Looks like a honeycomb.

Of course the dual space $(\mathbb{R}^n)^*$ is isomorphic to $\mathbb{R}^n$:

$$f = f_i \omega^i = f_i x^i \in (\mathbb{R}^n)^* \leftrightarrow (f_i) = (f_1, \ldots, f_n) \in \mathbb{R}^n$$

and as a vector space itself, covector addition is just the usual parallelogram vector addition. The above hyperplane interpretation of the dual space covector addition occurs on the original vector space!

These same pictures apply to any finite dimensional vector space. The directed line segment / directed hyperplane pair difference in geometrical interpretation is one reason for carefully distinguishing V from V^* by switching index positioning.

For $\mathbb{R}^n$ the distinction between n-tuples of numbers which are vectors and covectors is made using matrix notation. Vectors in $\mathbb{R}^n$ are identified with column matrices and those in the dual space with row matrices

$$\left.\begin{aligned} u &= (u^1, \dots, u^n) \longleftrightarrow \begin{pmatrix} u^1 \\ \vdots \\ u^n \end{pmatrix} \\ f &= f_i \omega^i \Leftrightarrow \underbrace{(f_1, \dots, f_n)}_{} \Leftrightarrow (f_1 \cdots f_n) \end{aligned}\right\} \quad f(u) = f_i u^i = (f_1 \cdots f_n)\begin{pmatrix} u^1 \\ \vdots \\ u^n \end{pmatrix}$$

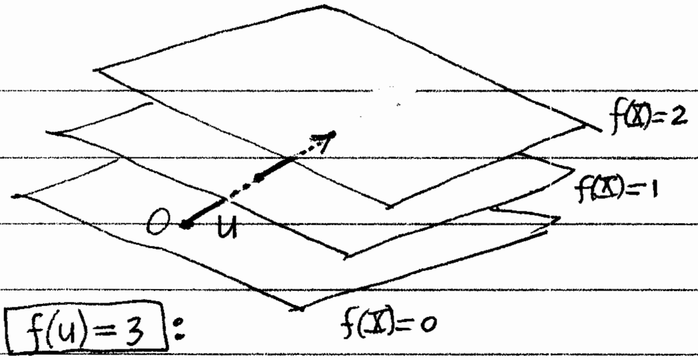
$$\equiv f_{(e)} \equiv (f_{(e)}{}^1, \dots, f_{(e)}{}^n)$$

subscript (e) since it depends on the basis $\{e_i\}$.

The evaluation of a covector on a vector (just the value of the function $f_i x^i$ at the point with position vector u) is a matrix product of two different objects, although it can be represented in terms of the usual dot product on $\mathbb{R}^n$ of two vectors (like objects)

$$f(u) = f_{(e)} \cdot u \equiv \sum_{i=1}^{n} f_{(e)}{}^i u^i$$

but the relationship between the covector f and the vector $f_{(e)}$ involves additional mathematical structure, that of an inner product on $\mathbb{R}^n$, which is associated with the Euclidean geometry we all know and love.



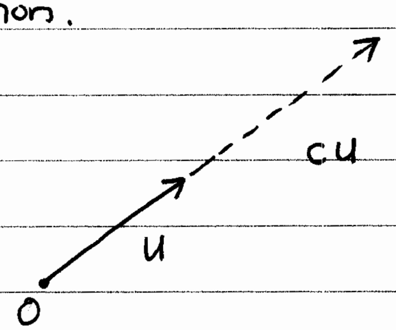
$\boxed{f(u)=3}$:

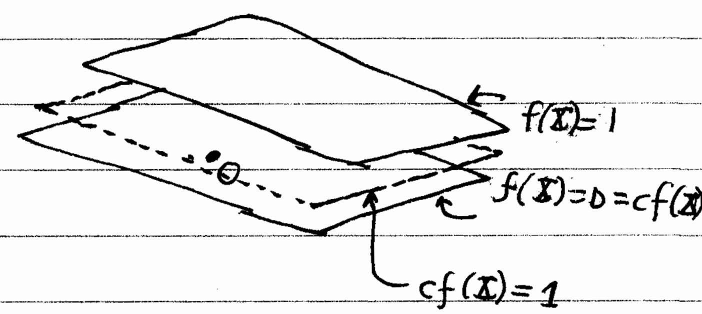
The evaluation of a covector on a vector also has a geometrical interpretation. If we imagine the 1-parameter family of hyperplanes $f_i x^i = f(X) = t$, then the value t of the evaluation is the "number" of hyperplanes pierced by the arrow representing u, if by "number" we refer to the integer subfamily and interpolate between them.

It is exactly this natural evaluation operation which is embodied in the Einstein index convention for summed repeated sub/super index pairs, which is different from a dot product $\quad f_{(e)} \cdot u = \sum_{i=1}^{n} f_{(e)}{}^i u^i$

which requires more structure as we will see below.

Scalar multiplication of vectors and covectors also has a geometrical interpretation.

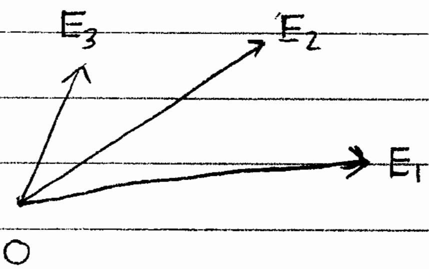




A vectors length is multiplied by $|c|$, with a direction reversal if $c<0$, but the separation between the two parallel planes representing a covector is <u>divided</u> by $|c|$ (with a direction reversal if $c<0$ in the sense that the plane $cf(\mathbf{x})=1$ is on the opposite side of the plane $f(\mathbf{x})=0$ from $f(\mathbf{x})=1$).

This increases the "number" of planes pierced by a given vector (assuming $|c|>1$), thus increasing the value of the covector on the vector.

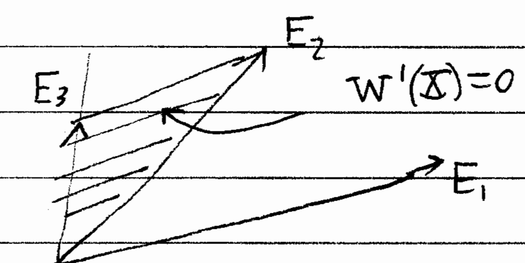
Using these geometrical pictures we can give a geometric construction of the dual basis to a given basis. Suppose we have 3 linearly independent vectors in $\mathbb{R}^3$. They form a basis. What is the parallel plane representation of the three dual basis vectors?



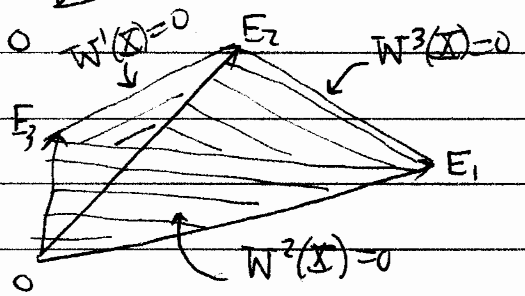
The dual basis is defined by the duality relations;

$$\begin{cases} W^i(E_j) = 0 & i \neq j \\ W^i(E_i) = 1 & (\text{no sum on } i) \end{cases}$$

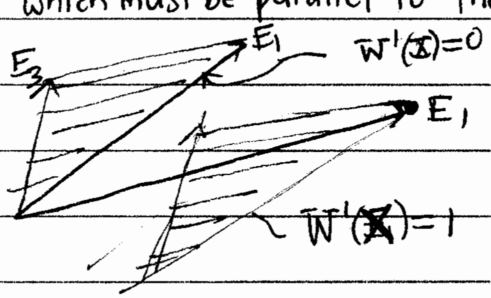
The first (offdiagonal) relation says that a given dual basis vector W^i should "kill" (give zero on) the "other" $(j \neq i)$ basis vectors, and hence on any linear combination of the other vectors and hence on any vector in the plane (hyperplane in $\mathbb{R}^n$) spanned by the other vectors.



So the plane of E_2 and E_3 is the plane $W^1(X)=0$ for W^1, and similarly for the others. So weve determined the orientations of each of the dual basis covectors from the "offdiagonal" duality relations. The "magnitude" is determined by the "diagonal" relations.



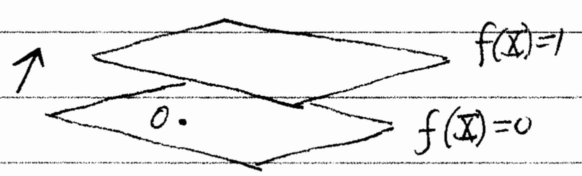
The relation $W^1(E_1) = 1$ means that the tip of E_1 lies in the plane $W^1(X)=1$ which must be parallel to the plane $W^1(X)=0$, completely fixing the former.



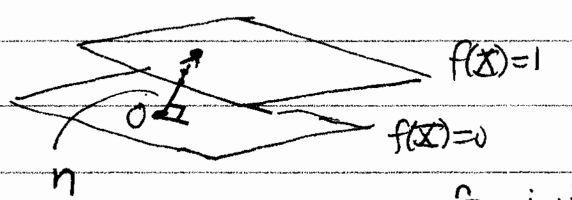
Drawing in all 3 pairs of planes makes a 3-dimensional honeycomb structure which I wont attempt to draw.

Notice that changing E_1 for fixed E_2, E_3 does not change the orientation of W^1, only the "magnitude" or the separation parameter. On the other hand changing E_1 doesn't affect the magnitude of W^2, W^3 but does change their orientation. These "complementary" effects of such a change reflect the "duality" between vectors and covectors.

A vector is represented by a directed line segment (1-dimensional) while a covector is represented by a directed pair of parallel planes of dimension $3-1=2$ in $\mathbb{R}^3$ or hyperplanes (dimension $n-1$) in $\mathbb{R}^n$.



The "directed" qualifier refers to the sense in which we start at the 0-value and finish at the 1-value, although there is no particular direction from the origin (i.e. specific vector) along which we go, unless we introduce a Euclidean geometry, for example, which picks out the direction orthogonal to the pair of planes. Suppose n is the orthogonal vector in $\mathbb{R}^3$ from the origin to the second plane.



Then $f(n) = f_i n^i = 1$, but if $f(X) = f_i X^i = 0$, then by orthogonality

$$0 = n \cdot X = \sum_{i=1}^{n} n^i X^i \equiv n_i X^i \quad (\text{let } n_i = n^i \text{ for each } i$$

to make indices come out right). Thus f and the covector (n_i) have the same level-surface with value 0 and so much be proportional, but it will turn out later that in fact the Euclidean lengths of their component vectors are inverses of one another. because of this relation

More later. For now it suffices to say that we can't pick out a particular direction in which the level surfaces of a covector increase in their value without additional structure.

Well, we began with a vector space V and introduced its dual space V^* which is a vector space in its own right and so has its own dual space $(V^*)^* \equiv V^{**}$ of real-valued linear functions of covectors:

$$F \in V^{**} \quad \text{means} \quad F(af+bg) = aF(f)+bF(g) \quad \textit{(linearity condition)}$$

However, unlike a relationship between a vectorspace and its dual which requires additional mathematical structure, there is a "natural" identification of a vector space and the dual of its dual. Of course they are all n-dimensional vector spaces and therefore isomorphic, but one has to choose a basis to establish a particular isomorphism which then depends on that choice of basis ("unnatural", not independent of choice of basis).

For each $v \in V$, define a $\tilde{v} \in V^{**}$ by $\tilde{v}(f) = f(v)$ for any covector f. Then

$$\begin{aligned}
(a\tilde{u}+b\tilde{v})(f) &= a\tilde{u}(f)+b\tilde{v}(f) && \text{def. of lin. comb of functions}\\
&= af(u)+bf(v) && \text{def. of tilde}\\
&= f(au+bv) && \text{linearity of covector}\\
&= \widetilde{(au+bv)}(f) && \text{def of tilde.}
\end{aligned}$$

Thus the map $\sim : V \to V^{**}$ is a linear map since $\widetilde{(au+bv)} = a\tilde{u}+b\tilde{v}$ (since f arbitrary). It is also 1-1 since if $\tilde{u} = \tilde{v}$, then $\tilde{u}(f) = \tilde{v}(f)$ so $f(u) = f(v)$ and $f(u-v) = 0$ (linearity) for every covector f, which can only be true if $u-v=0$ or $u=v$. This means it is a vector space isomorphism (1-1 linear map).

So if you start with $F \in V^{**}$, then

$$F(f) = F(f_i\omega^i) = f_i \underbrace{F(\omega^i)}_{\equiv F^i} = f_i F^i = f_i \delta^i{}_j F^j = f_i \omega^i(e_j) F^j$$

"components wrt $\{e_i\}$ of F"

$$= (f_i\omega^i)(F^j e_j) = f(F^j e_j),$$

i.e., evaluation of F on f is equivalent to evaluation of f on the vector $F^i e_i$. [$\{\tilde{e}_i\}$ is the basis dual to the basis $\{\omega^i\}$ of V^*, since $\tilde{e}_i(\omega^j) = \omega^j(e_i) = \delta^j{}_i$.]
We can therefore forget about V^{**} by using evaluation of covectors on vectors to produce linear functions of covectors.
Thus the natural evaluation $\quad f(v) = f_i v^i$
can be interpreted as a linear function of v for fixed f or as a linear function of f for fixed v. This puts vectors and covectors on an equal footing with respect to evaluation, and sometimes this is made explicit by using the notation

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

"scalar product" of covector and vector

which eliminates having to write one evaluated on the other as function notation requires.

[Another indication of natural/unnatural isomorphisms is that each time we go to the dual space we interchange index positions — after two interchanges they are back in the right position so one doesn't need additional structure to get the indices back in the "right position" as is necessary in the relationship between a vector space and its own dual: $F = F^i \tilde{e}_i \leftrightarrow F^i e_i$]

Linear transformations of a vector space into itself

Suppose $A : V \to V$ is a linear transformation, i.e., a V-valued linear function on V, or a linear function on V with values in V.

By linearity

$$A(v) = A(v^i e_i) = v^i \underbrace{A(e_i)}_{\equiv A^j{}_i e_j} = (A^j{}_i v^i)\, e_j$$

for each i, result $A(e_i)$ is a vector with components $A^j{}_i \equiv \omega^j(A(e_i))$

note natural index positions up/down

or $[A(v)]^j = A^j{}_i v^i$.

The jth component of the image vector is the jth entry of the matrix product of the matrix $\underline{A} \equiv (A^j{}_i)$ [$j \sim$ rows, $i \sim$ columns] [underline like in boldface notation for matrices] with the column vector $\underline{v} \equiv (v^i)$.

The matrix $\underline{A} = (\omega^j(A(e_i)))$ is the "matrix of A wrt the basis $\{e_i\}$." Obviously if you change the basis, the matrix will change. We'll get to that later.

Even if we are not working with $\mathbb{R}^n$, any choice of basis $\{e_i\}$ of V establishes an isomorphism with $\mathbb{R}^n$ (n-tuplet of components of a vector wrt this basis is a point in $\mathbb{R}^n$ — this essentially identifies the basis $\{e_i\}$ of V with the standard basis of $\mathbb{R}^n$).

Expressing everything in components leads us to matrix notation. Vectors in component form become column matrices, covectors become row matrices, and the linear transformation becomes a square matrix acting by matrix multiplication on the

$$([A(v)]_1 \cdots [A(v)]_n) = \underline{A}\ \underline{v}$$

$$f(v) = \underline{f}\ \underline{v}$$

$$\underline{A} = (A^i{}_j)$$

$$\underline{v} = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$$

$$\underline{f} = (f_1 \cdots f_n)$$

For every linear transformation A, we can define an associated bi-linear real-valued function of a pair of arguments consisting of a covector and a vector. [Bi-linear means linear in each of two arguments.]

$$\mathcal{A}(f,v) \equiv \underbrace{f(\underbrace{A(v)}_{\text{vector}})}_{\text{real number}} = (f_i\omega^i)(A^j{}_k v^k e_j) = f_i A^j{}_k v^k \omega^i(e_j)$$
$$= f_i A^j{}_k v^k \delta^i{}_j = f_i A^i{}_k v^k$$

For fixed f, this is a real-valued linear function of v, namely the covector with components $f_i A^i{}_k$ (one free down index).

For fixed v, this is a real-valued linear function of f, namely evaluation on the vector with components $A^i{}_k v^k$ (one free up index).

In general a $\binom{p}{q}$-tensor over V will simply be a real-valued multilinear function of p covector arguments (listed first) and q vector arguments (listed last):

$$T(\underbrace{f, g, \ldots}_{p}, \underbrace{v, u, \ldots}_{q}) \in \mathbb{R}.$$

By definition then, a covector is a $\binom{0}{1}$-tensor over V (1 vector argument, no covector arguments) while a vector is a $\binom{1}{0}$-tensor over V (1 covector argument, no vector argument) recalling that

$$v(f) \equiv f(v) \quad \text{(the value of a vector on a covector is the value of the covector on the vector).}$$

Thus a linear transformation A has (naturally) a $\binom{1}{1}$-tensor over V associated with it. Anytime we have a space of linear functions over a vector space [in this case the Cartesian product of pairs (f,v) of covectors and vectors], it has a natural linear structure by defining linear combinations of functions through linear combination of values, i.e., is itself a vector space & we can look for a basis.

Let $V \otimes V^*$ denote the space of $\binom{1}{1}$-tensors over V. The symbol $\otimes$ is called the tensor product, explained below. The zero element of this vector space is the multilinear function:

$$\underset{\substack{\uparrow \\ \text{zero tensor}}}{\mathbb{O}}(f,v) = \underset{\substack{\uparrow \\ \text{zero real number}}}{0} \quad \leftrightarrow \quad \underset{\substack{\uparrow \\ \text{zero matrix}}}{O^i{}_j} = \omega^i(\underset{\substack{\uparrow \\ \text{zero linear transformation}}}{O}(e_j)) = 0$$

$O(e_j) = 0$ zero vector

whose square matrix of components is the zero matrix.

Another element in this space is the evaluation tensor associated with the identity transformation $Id(v) = v$

$$EVAL(f,v) = f(v) = f_i \delta^i{}_j v^j$$

$$\leftrightarrow \quad (EVAL)^i{}_j = \omega^i(Id(e_j)) = \omega^i(e_j) = \delta^i{}_j$$

whose square matrix of components is the unit matrix.
EVAL is sometimes called the unit tensor.

To come up with a basis we need a definition. Given a covector and a vector we can produce a $\binom{1}{1}$-tensor by the definition

$$[v \otimes f](g,u) \equiv g(v)\, f(u) = g_i \underbrace{(v^i f_j)}_{} u^j$$

matrix of components of $v \otimes f$:

$$(v \otimes f)(\omega^i, e_j) = \omega^i(v)\, f(e_j) = v^i f_j$$

The symbol $\otimes$ is called the tensor product and only serves to hold v and f apart until they acquire arguments to be evaluated on.

[It creates an object taking 2 arguments from two objects taking single arguments.]

The component expression shows that $v \otimes f$ is clearly bilinear in its arguments g and u, so it is a $\binom{1}{1}$-tensor.

We can use $\otimes$ to create a basis for $V \otimes V^*$ from a basis $\{e_i\}$ and its dual basis $\{\omega^i\}$, namely the set $\{e_j \otimes \omega^i\}$ of $n^2 = n \times n$ such tensors.

By definition $(e_j \otimes \omega^i)(g,u) = g(e_j)\,\omega^i(u) = g_j u^i = u^i g_j$.
We can use this to show the two conditions that they form a basis:
1) spanning set: $A(f,v) = \cdots = f_j A^j{}_k v^k = A^j{}_k \underbrace{v^k f_j}_{(e_j \otimes \omega^k)(f,v)}$

$$= (A^j{}_k\, e_j \otimes \omega^k)(f,v)$$

so $A = A^j{}_k\, e_j \otimes \omega^k$ since the two functions have the same values on all pairs of arguments. The components of A with respect to this basis are just the components of A with respect to $\{e_i\}$ introduced above: $A^j{}_k = A(\omega^j, e_k)$.

2) linear independence: if $A^j{}_k\, e_j \otimes \omega^k = 0 \leftarrow$ (zero tensor)
then evaluating both sides on the argument pair (ω^m, e_n) leads to

$$(A^j{}_k\, e_j \otimes \omega^k)(\omega^m, e_n) = 0(\omega^m, e_n) = 0$$

$$= A^j{}_k \underbrace{\omega^m(e_j)}_{\delta^m{}_j} \underbrace{\omega^k(e_n)}_{\delta^k{}_n} = A^m{}_n$$

so all the coefficients must be zero, proving linear independence.

Thus $V \otimes V^*$ is the space of linear combinations of tensor products of vectors with covectors, explaining the notation.

So we've taken the linear algebra of $\mathbb{R}^n$, as embodied in column matrices (vectors), row matrices (covectors), both with n entries, and square $n \times n$ matrices ($\binom{1}{1}$-tensors), and generalized them into the mathematical structure of a vector space V, its dual space V^* and their tensor product space $V \otimes V^*$.
This abstracts from the 1 and 2 index objects associated with the elementary linear algebra of introductory courses to allow us to

consider objects in an invariant way (no indices) that correspond to any number of indices on those objects. Clearly we can play the same game with any space of tensors over V with arbitrary numbers of arguments of either type.

$$T = \binom{p}{q}\text{-tensor over } V: \qquad T(f, g, \ldots, V, U, \ldots) \in \mathbb{R}$$

Define its components with respect to $\{e_i\}$ by

$$T^{ij\ldots}{}_{mn\ldots} = T(\omega^i, \omega^j, \ldots, e_m, e_n, \ldots) \qquad \text{(real numbers)}$$

Introduce the basis $\{\underbrace{e_i \otimes e_j \otimes \cdots}_{p} \otimes \underbrace{\omega^m \otimes \omega^n \otimes \cdots}_{q}\} \leftarrow n^{p+q}$ basis "vectors" i.e. $\binom{p}{q}$-tensors

where $(f \otimes g \otimes u \otimes v)(w, z, h, i) = f(w)\, g(z)\, h(u)\, v(i)$

etc defines any number of factors in a tensor product

and make the expansion

$$T = T^{ij\ldots}{}_{mn\ldots}\, e_i \otimes e_j \otimes \cdots \otimes \omega^m \otimes \omega^n \otimes \cdots$$

EX. On $\mathbb{R}^3$ introduce the $\binom{0}{3}$-tensor D by

$$D(u, v, w) = u \cdot (v \times w) = \det \begin{pmatrix} u^1 & u^2 & u^3 \\ v^1 & v^2 & v^3 \\ w^1 & w^2 & w^3 \end{pmatrix} \qquad \text{"triple scalar product"}$$

This is linear in each vector argument (the determinant is a linear function of each row). It therefore has the expansion

$$D = D_{ijk}\, \omega^i \otimes \omega^j \otimes \omega^k \qquad \text{where}$$

$$D_{ijk} = D(e_i, e_j, e_k) = \begin{cases} 1 \text{ if } (i,j,k) \text{ even permutation of } (1,2,3) \\ -1 \text{ if } (i,j,k) \text{ odd permutation of } 1,2,3 \\ 0 \text{ otherwise} \end{cases}$$

$$D = \omega^1 \otimes \omega^2 \otimes \omega^3 + \omega^2 \otimes \omega^3 \otimes \omega^1 + \omega^3 \otimes \omega^2 \otimes \omega^1 - \omega^1 \otimes \omega^3 \otimes \omega^2 - \omega^2 \otimes \omega^1 \otimes \omega^3 - \omega^3 \otimes \omega^1 \otimes \omega^2 .$$

Still on $\mathbb{R}^3$ introduce the $\binom{0}{4}$-tensor

$$C(u,v,w,z) = (u\times v)\cdot(w\times z),$$ "quadruple scalar product"

or the $\binom{0}{2}$ tensor G defined by "dot product"

$$G(u,v) = u\cdot v = \delta_{ij}u^i v^j \quad \text{where } \delta_{ij} \doteq \begin{cases} 1 & i=j \\ 0 & i\neq j. \end{cases}$$

To make the index position work we need to introduce a Kronecker delta with both indices down. which we interpret as the components $\delta_{ij} = G(e_i, e_j) = e_i \cdot e_j$ of the $\binom{0}{2}$-tensor G, namely the dot products of the basis vectors, so

$$G = \delta_{ij}\,\omega^i\otimes\omega^j = \omega^1\otimes\omega^1 + \omega^2\otimes\omega^2 + \omega^3\otimes\omega^3.$$

An inner product on any vector space is a "symmetric" $\binom{0}{2}$-tensor which accepts two vector arguments *in either order* and produces a real number. [and such that the determinant of its symmetric matrix of components is nonzero] The dot product on $\mathbb{R}^n$ is such an inner product whose matrix of components is the identity matrix with respect to the standard basis of $\mathbb{R}^n$

The index positioning δ_{ij} for a $\binom{0}{2}$ tensor shows that it is fundamentally different from the identity $\binom{1}{1}$-tensor with components $\delta^i{}_j$, even though both matrices of components are the unit matrix. More of this later.

The simplest example of a dot product created tensor is a covector:

$$f_u(v) = u \cdot v$$

For each fixed u, this defines a linear function of v, i.e., a covector f_u. It is exactly this correspondence that allows one to avoid covectors in elementary linear algebra.

Tensor product and matrix multiplication

By linearity, the components of the tensor product of a vector and a covector are

$$v\otimes f = (v^i e_i)\otimes(f_j\omega^j) = \underbrace{v^i f_j}_{(v\otimes f)^i{}_j = v^i f_j}\, e_i\otimes\omega^j$$

or equivalently $(v\otimes f)^i{}_j = (v\otimes f)(\omega^i, e_j) = \omega^i(v)\,f(e_j) = v^i f_j$.

With the representation in component form of a vector and a covector as column and row matrices respectively, this tensor product is exactly equivalent to matrix multiplication

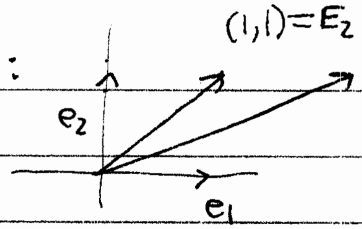
$$\underset{\substack{\uparrow \\ \text{col matrix}}}{\underline{v}}\ \underset{\substack{\uparrow \\ \text{row matrix}}}{\underline{f}} = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}(f_1 \cdots f_n) = \begin{pmatrix} v^1 f_1 & \cdots & v^1 f_n \\ \vdots & & \vdots \\ v^n f_1 & \cdots & v^n f_n \end{pmatrix} = (v^i f_j),$$

but in the opposite order from the evaluation of a covector on a vector, leading to a matrix rather than a scalar (number).

Thus matrix multiplication of a row matrix by a column matrix on the right represents the abstract evaluation operation of a covector on a vector or viceversa, while the matrix multiplication on the left represents the tensor product operation. In this sense the name "scalar product" for evaluation is more analogous to tensor product (one produces a "scalar," or real number, the other a tensor).

PROBLEM:

$\mathbb{R}^2$



a) Find the dual basis $\{W^1, W^2\}$ to $\{E_1, E_2\}$ in terms of $\{\omega^1, \omega^2\}$.

b) What is $W^2((5,-2))$?

c) Plot the vectors (A,B) and (C,D) of the hint on the same axes. What do you notice about their relation to $\{E_i\}$?

HINT: $W^1 = A\omega^1 + B\omega^2$ (why?)

$W^2 = C\omega^1 + D\omega^2$

so express 4 equations $W^i(E_j) = 0$ to obtain 4 equations to determine the 4 constants A, B, C, D.

ANSWER:

$$W^1(E_1) = A\omega^1(E_1) + B\omega^2(E_1) = A\cdot 2 + B\cdot 1 = \delta^1{}_1 = 1$$
$$W^1(E_2) = A\omega^1(E_2) + B\omega^2(E_2) = A\cdot 1 + B\cdot 1 = \delta^1{}_2 = 0$$
$$W^2(E_1) = C\omega^1(E_1) + D\omega^2(E_1) = C\cdot 2 + D\cdot 1 = \delta^2{}_1 = 0$$
$$W^2(E_2) = C\omega^1(E_2) + D\omega^2(E_2) = C\cdot 1 + D\cdot 1 = \delta^2{}_2 = 1.$$

solving these two simple systems givens $(A,B) = (1,-1)$, $(C,D) = (-1,2)$.

In fact $W^1(u) = (A,B)\cdot u$ so if $W^1(u) = 0$, the u is orthogonal to the vector (A,B) of components of W^1 wrt $\{e_i\}$. Thus $\vec{W}^1 \equiv (1,-1)$ is orthogonal to E_2 and $\vec{W}^2 \equiv (-1,2)$ is orthogonal to E_1.

$$W^2((5,-2)) = (-1,2)\cdot(5,-2) = -5-4 = -9$$
or
$$= -\omega^1((5,-2)) + 2\omega^2((5,-2)) = -5 + 2(-2) = -9.$$

We could have written $W^i = W^i{}_j\,\omega^j$, leading to a matrix

ith covector ↑ jth component wrt $\{e_i\}$ of W^i

$W^i{}_j = W^i(e_j)$ which "changes the basis". More later.

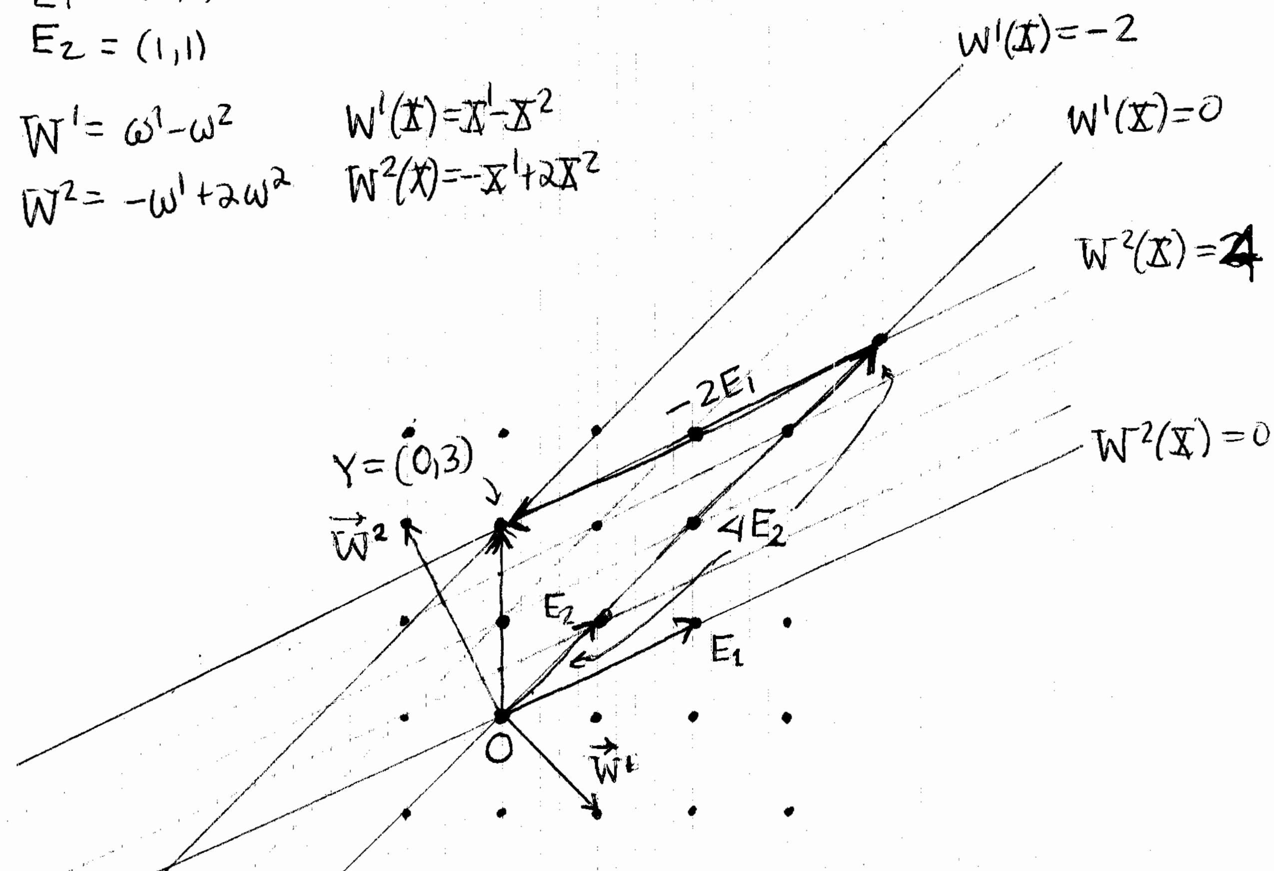
new old

$\{E_i\}$, dual basis $\{W^i\}$ on $\mathbb{R}^2$

$E_1 = (2,1)$
$E_2 = (1,1)$

$W^1 = \omega^1 - \omega^2 \qquad W^1(X) = X^1 - X^2$

$W^2 = -\omega^1 + 2\omega^2 \qquad W^2(X) = -X^1 + 2X^2$



$$Y = W^1(Y)E_1 + W^2(Y)E_2$$
$$= -2E_1 + 4E_2$$

Here is a graphical representation of the integer level surfaces of W^1 and W^2 and an example of decomposing a vector into components with respect to $\{E_i\}$ using the dual basis.

Note: $W^1 = x - y$, $W^2 = -x + 2y$ in cartesian coordinates $\{x, y\}$ on $\mathbb{R}^2$.

PROBLEM a) Find the $\binom{1}{1}$-tensor A expressed in terms of $\{e_i\}$ and $\{\omega^i\}$ using the previous problem on $\mathbb{R}^2$ with new basis $\{E_1, E_2\} = \{(2,1), (1,1)\}$ and dual basis $W^1 = \omega^1 - \omega^2$, $W^2 = -\omega^1 + 2\omega^2$, if A has the following components in terms of the basis $\{E_i\}$:

$$A(W^1, E_1) = 1 \qquad A(W^1, E_2) = 2$$
$$A(W^2, E_1) = -1 \qquad A(W^2, E_2) = 0.$$

b) what is its matrix of components with respect to $\{e_i\}$?

ANSWER If we let $A^i{}_j = A(W^i, E_j)$ be the components of A with respect to $\{E_i\}$, then

$$A = A^i{}_j \, E_i \otimes W^j = E_1 \otimes W^1 + 2E_1 \otimes W^2 - E_2 \otimes W^1.$$

But both $\{E_i\}$ and $\{W^j\}$ are linear combinations of the standard basis and dual basis, so we can just substitute and expand.

$$A = \underbrace{(2e_1+e_2)\otimes(\omega^1-\omega^2) + 2(2e_1+e_2)\otimes(-\omega^1+2\omega^2)}_{(2e_1+e_2)\otimes[\underbrace{(\omega^1-\omega^2)+2(-\omega^1+2\omega^2)}_{-\omega^1+3\omega^2}]} - (e_1+e_2)\otimes(\omega^1-\omega^2)$$

$$= -2e_1\otimes\omega^1 - e_2\otimes\omega^1 + 6e_1\otimes\omega^2 + 3e_2\otimes\omega^2 - e_1\otimes\omega^1 - e_2\otimes\omega^1 + e_1\otimes\omega^2 + e_2\otimes\omega^2$$

$$= \underbrace{-3}_{A(\omega^1,e_1)} e_1\otimes\omega^1 + \underbrace{7}_{A(\omega^1,e_2)} e_1\otimes\omega^2 \underbrace{-2}_{A(\omega^2,e_1)} e_2\otimes\omega^1 + \underbrace{4}_{A(\omega^2,e_2)} e_2\otimes\omega^2 = A(\omega^i, e_j)\, e_i \otimes \omega^j$$

These are the components with respect to $\{e_i\}$. The matrix is

$$\left(A(\omega^i, e_j)\right) = \begin{pmatrix} -3 & 7 \\ -2 & 4 \end{pmatrix}$$ assuming my arithmetic is correct.

Remark Our notation is so compact that certain facts may escape us.
For example $v \otimes f = (v^i e_i) \otimes f = v^i e_i \otimes f$ is actually a distributive law for the tensor product. A simpler example shows this

$$(u+v) \otimes f = u \otimes f + v \otimes f.$$

How do we know this? Well, the only thing we know about the tensor product is how it is defined in terms of evaluation on its arguments

$$
\begin{aligned}
[(u+v) \otimes f](g, w) &\equiv g(u+v) f(w) = [g(u) + g(v)] f(w) \\
&= g(u) f(w) + g(v) f(w) \\
&= (u \otimes f)(g, w) + (v \otimes f)(g, w) \\
&= [u \otimes f + v \otimes f](g, w) \quad \text{"how one adds functions to produce the sum function"}
\end{aligned}
$$

But if these functions inside the square brackets on each side of the equation have the same values on all pairs of arguments, they are the same function (i.e. $\binom{1}{1}$-tensor)

$$(u+v) \otimes f = u \otimes f + v \otimes f.$$

In fact it is easy to show (exercise) that $(cv) \otimes f = c(v \otimes f)$ for any constant c, so in fact the tensor product behaves like a product should with linear combinations.