Linear transformations and a change of basis

Suppose $A : V \to V$ is a linear transformation of V into itself.
If $\{e_i\}$ is a basis of V, then the matrix of A wrt $\{e_i\}$
is defined by (page 10)

$$\underline{A} = (A^i{}_j) \quad , \qquad A^i{}_j = \omega^i(A(e_j)) = i\text{th component of } A(e_j).$$

i is the row index and j the column index (first and second indices respectively, although the first is a superscript instead of the usual subscript like the second).
The jth column is the column matrix of components of $A(e_j)$ wrt the basis, denoted by underlining

$$\underline{A} = \left(\underline{A(e_1)} \ \underline{A(e_2)} \cdots \underline{A(e_n)} \right).$$

If we expand the equation

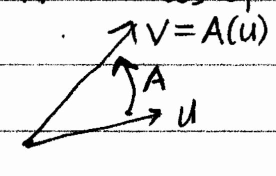
$$u = A(v)$$

$$u^i e_i \qquad \qquad A(v^j e_j) = v^j A(e_j) = v^j A^i{}_j e_i = (A^i{}_j v^j) e_i$$

we get the component relation $\quad u^i = A^i{}_j v^j$

or its matrix form $\quad \underline{u} = \underline{A}\, \underline{v}$

where $\underline{u} = \begin{pmatrix} u^1 \\ \vdots \\ u^n \end{pmatrix}, \underline{v} = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$ are the column matrices of components of u and v wrt the basis.



We can interpret this as an "active" transformation of the points (vectors) of V to new points of V. We start with a vector $\mathbf{v}$ and end up at the new vector $\mathbf{u}$.

We can also use a linear transformation to change the basis of V, provided that it is nonsingular (its matrix has nonzero determinant), just the condition that the n image vectors of the original basis $\{e_i\}$ are linearly independent and so can be used as a new basis.

If $B: V \to V$ is such a linear transformation, with matrix $\underline{B} = (B^i{}_j) = \omega^i(B(e_j))$, $\det \underline{B} \neq 0$, then define $e_{i'} = B(e_i) = B^j{}_i e_j$.

As discussed above, the columns of $\underline{B} = (\underline{B(e_1)} \cdots \underline{B(e_n)})$ are the components of the new basis vectors with respect to the old ones $B^i{}_j = \omega^i(B(e_j)) \equiv \omega^i(e_{j'})$.

Primed indices will be associated with component expressions in the new basis.

Since B is invertible, we have

$$e_i = B^{-1}(e_{i'}) = B^{-1\,j}{}_i\, e_{j'}.$$

The new basis $\{e_{i'}\}$ has its own dual basis $\{\omega^{i'}\}$ satisfying

$$\omega^{i'}(e_{j'}) = \delta^i{}_j$$

If we define $\omega^{i'} = B^{-1\,i}{}_j\, \omega^j$, then

$$\omega^{i'}(e_{j'}) = B^{-1\,i}{}_k\, \omega^k(B^\ell{}_j e_\ell) = B^{-1\,i}{}_k B^\ell{}_j \delta^k{}_\ell$$

$$= B^{-1\,i}{}_k B^k{}_j = \delta^i{}_j \qquad \text{since } \underline{B}^{-1}\underline{B} = \underline{I},$$

so this is the correct expression for the new dual basis.

Given any vector v, we can express it either in terms of the old basis or the new

$$v = \begin{cases} v^i e_i & v^i = \omega^i(v) \\ v^{i'} e_{i'} & v^{i'} = \omega^{i'}(v) = B^{-1\,i}{}_j\, \omega^j(v) = B^{-1\,i}{}_j\, v^j \end{cases}$$

In other words, if we actively transform the old basis to a new basis using B, the new components of any vector are related to the old components of the same vector by matrix multiplication by the inverse matrix $\underline{B}^{-1}$:

$$\underline{V}' = \underline{B}^{-1}\underline{V} \quad \text{or equivalently} \quad \underline{V} = \underline{B}\,\underline{V}'.$$

Similarly we can express any covector in terms of the old or new dual basis

$$f = f_i \omega^i \qquad f_i = f(e_i)$$

$$= f_{i'}\omega^{i'} \qquad f_{i'} = f(e_{i'}) = f(B^j{}_i e_j) = B^j{}_i f(e_j) = f_j B^j{}_i \,,$$

i.e. the covector components transform by the matrix B but multiplying from the right if we represent covectors as row matrices

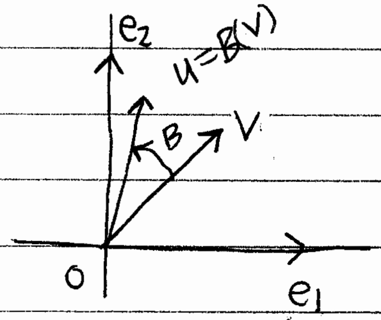
$$(f_{1'} \cdots f_{n'}) = (f_1 \cdots f_n)\,\underline{B}$$

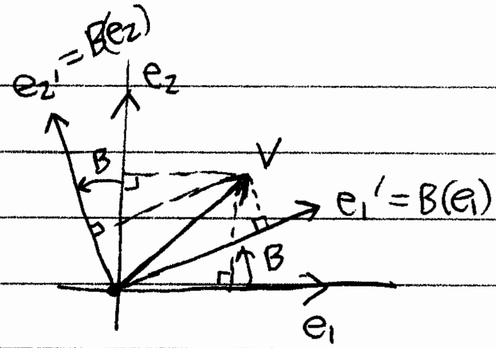
$$\underline{f}' = \underline{f}\,\underline{B} \qquad \text{or equivalently} \quad \underline{f} = \underline{f}'\underline{B},$$

where it is understood from the matrix context that $\underline{f}$ and $\underline{f}'$ are row vectors.

This describes a "passive" transformation of V into itself or of V^* into itself, since the points of these spaces do not change but their components do due to the change of basis. (doo doo?)

Changing the basis actively by a linear transformation B makes the components of vectors change by the inverse matrix $\underline{B}^{-1}$ of B, while an active transformation of V into itself gives the components with respect to the unchanged basis of the new vectors as the matrix product by $\underline{B}$ with the old components. The active and passive transformations go in opposite directions so to speak.





components: $\underline{v} = \underline{B}\,\underline{u}$ — active, basis fixed — vectors move

$\underline{v}' = \underline{B}^{-1}\underline{v}$ — passive, basis moves — vectors fixed.

If we are more interested in changing bases than active linear transformations, we can let $A = B^{-1}$. Then we have

$$\omega^{i'} = A^i{}_j\,\omega^j \quad\rightarrow\quad v^{i'} = A^i{}_j\,v^j$$

$$e_{i'} = A^{-1\,j}{}_i\,e_j \quad\rightarrow\quad f_{i'} = f_j\,A^{-1\,j}{}_i$$

Thus upper indices associated with vector component labels transform by the matrix $\underline{A}$, while lower indices associated with covector component labels transform by the matrix $\underline{A}^{-1}$.

In the jargon of this subject, these upper indices on components are called "contravariant" while the lower indices on components are called "covariant". Vectors and covectors themselves are sometimes called "contravariant vectors" and "covariant vectors" respectively. The above relations between old and new components of the same object are called "transformation laws" for contravariant and covariant vector components.

By the linearity of the tensor product, these "laws" can be extended to the components of any tensor.

For example, suppose $\mathbb{L} = L^i{}_j e_i \otimes \omega^j$ is the $\binom{1}{1}$- tensor associated with a linear transformation $L : V \to V$.

Then

$$\mathbb{L} = \begin{cases} L^i{}_j e_i \otimes \omega^j & L^i{}_j = \mathbb{L}(\omega^i, e_j) \\ L^{i'}{}_{j'} e_{i'} \otimes \omega^{j'} & L^{i'}{}_{j'} = \mathbb{L}(\omega^{i'}, e_{j'}) = \mathbb{L}(A^i{}_k \omega^k, A^{-1\ell}{}_j e_\ell) \end{cases}$$

$$= A^i{}_k A^{-1\ell}{}_j \underbrace{\mathbb{L}(\omega^k, e_\ell)}_{L^k{}_\ell} = A^i{}_k A^{-1\ell}{}_j L^k{}_\ell$$

In other words the contravariant and covariant indices each transform independently:

$$L^{i'}{}_{j'} = A^i{}_k A^{-1\ell}{}_{j'} L^k{}_\ell$$

or inversely:

$$L^i{}_j = A^{-1i}{}_k A^\ell{}_j L^{k'}{}_{\ell'} .$$

This generalizes in an obvious way to any $\binom{p}{q}$- tensor:

$$T = \begin{cases} T^{i\cdots}{}_{j\cdots} e_i \otimes \cdots \otimes \omega^j \otimes \cdots & T^{i\cdots}{}_{j\cdots} = T(\omega^i, \ldots, e_j, \ldots) \\ T^{i'\cdots}{}_{j'\cdots} e_{i'} \otimes \cdots \otimes \omega^{j'} \otimes \cdots & T^{i'\cdots}{}_{j'\cdots} = T(\omega^{i'}, \ldots, e_{j'}, \ldots) \end{cases}$$

$$= \cdots$$

$$= A^i{}_k \cdots A^{-1\ell}{}_j \cdots T^{k\cdots}{}_{\ell\cdots}$$

It is just a simple consequence of multilinearity.

EX. We first defined the Kronecker delta just as a convenient shorthand symbol $\delta^i{}_j$, but then saw that it coincided with the components of the evaluation or identity tensor

$$\text{EVAL} = \delta^i{}_j e_i \otimes \omega^j = e_i \otimes \omega^i = e_1 \otimes \omega^1 + \cdots + e_n \otimes \omega^n .$$

Since this must be true in any basis, if we "transform" the Kronecker delta as the components of a $\binom{1}{1}$-tensor, it should be left unchanged

$$\delta^{i'}{}_{j'} = A^i{}_k A^{-1\ell}{}_j \delta^k{}_\ell = A^i{}_k A^{-1k}{}_\ell = (\underline{A}\,\underline{A}^{-1})^i{}_\ell = \delta^i{}_\ell .$$

The new components do equal the old!

Matrix form of "transformation law" for $\binom{1}{1}$-tensors

The "transformation law" for the $\binom{1}{1}$-tensor $\mathbb{L}$ associated with a linear transformation $L: V \to V$ is:

$$L^{i'}{}_{j'} = A^{i'}{}_{k} A^{-1\ell}{}_{j'} L^{k}{}_{\ell} = \underbrace{A^{i'}{}_{k} L^{k}{}_{\ell}}\, \underbrace{\ }A^{-1\ell}{}_{j'}$$

$$= [\underline{A}\ \underline{L}\ \underline{A}^{-1}]^{i}{}_{j}$$

In other words we recover the matrix transformation for a linear transformation under a change of basis

$$\underline{L}' = \underline{A}\ \underline{L}\ \underline{A}^{-1}$$

which leads to the conjugation operation (sandwiching between a matrix and its inverse), except it was written in terms of the inverse $\underline{A}^{-1} = \underline{B}$:

$$\underline{L}' = \underline{B}^{-1}\underline{L}\ \underline{B} \longleftarrow \text{columns of } B = \text{old components of new basis vectors or}$$

$$\underline{v} = \underline{B}\,\underline{v}'$$

Ditto for $\binom{0}{2}$-tensors

A $\binom{0}{2}$-tensor has 2 vector arguments $\quad G(u,v) \in \mathbb{R}$, $\quad G = G_{ij}\,\omega^i \otimes \omega^j$.

The transformation law is

$$G_{i'j'} = A^{-1m}{}_{i} A^{-1n}{}_{j} G_{mn} = \underbrace{A^{-1m}{}_{i}\, G_{mn}}_{\text{not adjacent}}\, \underbrace{A^{-1n}{}_{j}}_{\text{adjacent}} = [(\underline{A}^{-1})^T \underline{G}\ \underline{A}]_{ij}$$

$$= [\underline{B}^T \underline{G}\,\underline{B}]_{ij}$$

Although G also has a matrix representation $\underline{G} = (G_{ij})$ in component form, its matrix transformation law involves the transpose, rather than the inverse. The transpose is necessary to represent to summation of the row index of $\underline{A}^{-1}$ against the row index of $\underline{G}$ in terms of matrix multiplication.

Recall that it was the transpose version of the law used in linear algebra diagonalization discussions. Of course if only orthogonal matrices $\underline{B}$ are used (restriction to orthonormal coordinates) $B^{-1} = B^T$ and there is no difference. But that requires the dot product to discuss and therefore additional structure.

$\underline{\text{EX}}$. The triple scalar product on $\mathbb{R}^3$ is a multilinear function on triplets of vectors, namely the determinant function, which is a $\binom{0}{3}$-tensor.

$$D(u,v,w) = u \cdot (v \times w) = \det \begin{pmatrix} u^1 u^2 u^3 \\ v^1 v^2 v^3 \\ w^1 w^2 w^3 \end{pmatrix}.$$

Suppose we define the shorthand symbol for it's components

$$\epsilon_{ijk} = D(e_i, e_j, e_k) = \begin{cases} 1 & \text{if } (i,j,k) = \text{posi perm. of } (1,2,3) \\ -1 & \text{if } (i,j,k) = \text{neg. perm. of } (1,2,3) \\ 0 & \text{otherwise} \end{cases}$$

Recall that $(1,2,3)\ (2,3,1)\ (3,1,2)$ are the positive permutations of $(1,2,3)$ while $(3,2,1), (2,1,3)\ (1,3,2)$ are the negative permutations.

Suppose we evaluate the new components of the determinant tensor

$$D = D_{ijk}\, \omega^i \otimes \omega^j \otimes \omega^k = D_{i'j'k'}\, \omega^{i'} \otimes \omega^{j'} \otimes \omega^{k'}$$

$$\begin{aligned}
D_{i'j'k'} &= A^{-1m}{}_i\, A^{-1n}{}_j\, A^{-1p}{}_k\, D_{mnp} \\
&= \epsilon_{mnp}\, A^{-1m}{}_i\, A^{-1n}{}_j\, A^{-1p}{}_k \\
&= \sum_{\substack{\text{all permutations } \sigma \\ \text{of } (1,2,3)}} (-1)^{\operatorname{sgn}\sigma}\, A^{-1\sigma(1)}{}_i\, A^{-1\sigma(2)}{}_j\, A^{-1\sigma(3)}{}_k \\
&= \begin{cases} \det \underline{A}^{-1} & \text{if } (i,j,k) = (1,2,3) \quad \text{by definition} \\ \det \underline{A}^{-1} & \text{if } (i,j,k) = \text{pos per of } (1,2,3) \\ -\det \underline{A}^{-1} & \text{if } (i,j,k) = \text{neg per of } (123) \\ 0 & \text{otherwise (repeated rows)} \end{cases} \quad \text{by row permutation rules} \\
&= (\det \underline{A}^{-1})\, \epsilon_{ijk}
\end{aligned}$$

In other words the new components of the determinant function differ from the old by the factor $\det \underline{A}^{-1}$. So the symbol ϵ_{ijk} $\underline{\text{does}}$ not define a tensor in the sense that its components in any basis have these same values, as does the Kronecker delta symbol.

"permutation symbol"

Another way of stating this is that this "alternating symbol" ("Levicivita indicator") ϵ_{ijk} defines a different tensor for each choice of basis

$$D_{(e)} = \epsilon_{ijk}\,\omega^i\otimes\omega^j\otimes\omega^k \neq D_{(e')} = \epsilon_{ijk}\,\omega^{i'}\otimes\omega^{j'}\otimes\omega^{k'}.$$

So the important lesson from this is, if we define an object with indices not by taking components of some tensor, then it is not necessarily a tensor — but may define a different tensor in each choice of basis.

Another way of handling this problem is to generalize the idea of a tensor (independent of the choice of basis) to a "tensor density" which is a family of tensors, one in each choice of basis, related by a more general transformation law which not only changes the components of the tensor but also changes the tensor itself by an overall factor of ~~some~~ one power of the determinant of the transformation.

Dividing through by the determinant factor gives

$$\underbrace{\epsilon_{ijk}}_{\text{new components of new tensor}} = \underbrace{(\det A^{-1})^{-1}}_{\text{changes to new tensor}}\ \underbrace{A^{-1\,m}{}_{i}\,A^{-1\,n}{}_{j}\,A^{-1\,n}{}_{k}\,\underbrace{\epsilon_{mnn}}_{\text{old components of old tensor}}}_{\text{new components of old tensor}}$$

(this power = "weight" of tensor density

The permutation symbol may be interpreted (by definition) as the components of a $\binom{0}{3}$-tensor density of weight -1.

Linear transformations between V and V^*

So far we've generalized vectors, covectors, and $\binom{1}{1}$-tensors from column matrices, row matrices, and the matrices of linear transformations, but have not considered $\binom{0}{2}$-tensors and $\binom{2}{0}$-tensors which are also associated with certain matrices, except for the matrix form of the transformation law for the $\binom{0}{2}$-tensors on page 24 b. The $\binom{0}{2}$-tensors and $\binom{2}{0}$-tensors may be interpreted as linear transformations between the vector space and its dual.

For example, suppose $\ell : V \to V^*$ is a linear map. Define the associated $\binom{0}{2}$-tensor $\boldsymbol{\ell} = \ell_{ij}\,\omega^i \otimes \omega^j$ by

$$\boldsymbol{\ell}(u,v) = \underbrace{\underbrace{\ell(v)}_{\text{covector}}(u)}_{\text{scalar}} = \ell_{ij}u^iv^j = \underbrace{(\ell_{ij}v^j)}_{[\ell(v)]_i \,=\, \ell_{ij}v^j}u^i ,$$

or $\underbrace{\ell(v)}_{\text{row matrix}} = (\overbrace{\underline{\ell}\ \underline{v}}^{\text{column matrix}})^{T}$ ← to make a row matrix from column matrix

where the components are

$$\ell_{ij} = \boldsymbol{\ell}(e_i,e_j) = \ell(e_j)(e_i) .$$

The linear map ℓ is realized by evaluating the second argument of the tensor $\boldsymbol{\ell}$ on a vector, so that a covector remains waiting for the first argument of the tensor.

In component form the transformation is again matrix multiplication but because both matrix indices are down, one is left with a covector, requiring an additional transpose in matrix form to yield a row matrix as agreed for representing covectors.

These linear transformations LOWER the index position.

In exactly the same way a linear map $\mathcal{R}: V^* \to V$ has an associated $\binom{2}{0}$-tensor $\mathcal{R} = r^{ij} e_i \otimes e_j$

$$\mathcal{R}(f,g) = \underbrace{f(\underbrace{\mathcal{R}(g)}_{\text{vector}})}_{\text{scalar}} = r^{ij} f_i g_j = f_i (\underbrace{r^{ij} g_j})$$

$$[\mathcal{R}(g)]^i = r^{ij} g_j$$

$$\text{or} \quad \underline{\mathcal{R}(g)} = \underline{r}\, \underline{g}^T$$

to make column vector for matrix multiplication

where $\quad r^{ij} = \mathcal{R}(\omega^i, \omega^j) = \omega^i(\mathcal{R}(\omega^j))$.

The linear map is realized by evaluating the second argument of the tensor, leaving a vector waiting for the second argument of the tensor.

These linear transformations RAISE the index position.

Invertible maps between V and V^*

The images of the basis or dual basis vectors under such maps are

$$\ell(e_i) = \ell_{ij}\omega^j$$

$$r(\omega^i) = r^{ij} e_j \,.$$

The condition that these maps be invertible is just that their corresponding matrices be invertible, i.e., have nonzero determinants so that their inverses exist $\quad \det(\ell_{ij}) \neq 0, \quad \det(r^{ij}) \neq 0$.

$$\to \quad \underline{\ell}^{-1} \equiv (\ell^{ij}) \quad , \underline{r}^{-1} \equiv (r_{ij}) \quad \text{exist and}$$

$$\ell^{ij}\ell_{jk} = \delta^i{}_k \quad , \quad r^{ij} r_{jk} = \delta^i{}_k$$

These are the matrices of linear maps $\ell^{-1}: V^* \to V$ and $r^{-1}: V \to V^*$. Such a pair, whether we start with ℓ or r (set $r = \ell^{-1}$ and $\ell = r^{-1}$ for example) establishes an isomorphism between the vector space and its dual.

Although one can use an arbitrary nonsingular matrix (ℓ_{ij}) and its inverse (ℓ^{ij}) to play this game, in practice two special kinds of such matrices are used, either symmetric or antisymmetric.

symmetric: $\ell_{ji} = \ell_{ij}$ or $\ell(v,u) = \ell(u,v)$

antisymmetric: $\ell_{ji} = -\ell_{ij}$ or $\ell(v,u) = -\ell(u,v)$.

The corresponding tensors are also called symmetric or antisymmetric, and in this context are said to define respectively an inner product (symmetric case) or a symplectic form (antisymmetric case) and both (usually the symmetric case) are called metrics.

EX. On $\mathbb{R}^n$ the dot product defines a particular inner product

$$G(u,v) = u \cdot v = \delta_{ij} u^i v^j$$

$$G_{ij} = \delta_{ij} = G(e_i, e_j) = e_i \cdot e_j .$$

Note that if you change from the standard basis to an arbitrary basis, the components of G will change

$$G_{i'j'} = (A^{-1})^m{}_i (A^{-1})^n{}_j \, \delta_{mn} . \qquad (= e_{i'} \cdot e_{j'})$$

Only if $(A^i{}_j)$ is an orthogonal matrix will this again equal δ_{ij}. In other words the symbol δ_{ij} defines a different tensor in each basis unless one restricts to only orthonormal bases.

In a nonorthonormal basis, $G_{i'i'} \neq 1$ breaks the normality (unit vector) condition, while $G_{i'j'} \neq 0$ $(i \neq j)$ breaks the orthogonality condition.

EX. On $\mathbb{R}^4$ let $\eta_{11} = \eta_{22} = \eta_{33} = -\eta_{44} = 1$ and $\eta_{ij} = 0 \ (i \neq j)$.
Then $G = \eta_{ij}\, \omega^i \otimes \omega^j$ defines an inner product, called the Minkowski metric. This determines the geometry of spacetime in special relativity, called Lorentzian as opposed to Euclidean because of the single minus sign.

Exercise: Given an inner product G on a vector space V we can always use the dot product notation by defining

$$G(u,v) \equiv u \cdot v$$

The self-dot-product contains two independent pieces of information: its sign (the "type": $+, -, 0$) and its magnitude. Define the magnitude or length of a vector and its sign (or "type") by

$$\|v\| = |v \cdot v|^{1/2} = |G(v,v)|^{1/2}$$

$$\operatorname{sgn} v = \operatorname{sgn}(v \cdot v) = \operatorname{sgn}(G(v,v)) \in \{+, 0, -\}$$

(i) In $\mathbb{R}^n$, [with usual dot product] what is the value of $\operatorname{sgn} v$ for any $v \neq 0$?
[This makes $\mathbb{R}^n$ Euclidean]

(ii) In $\mathbb{R}^n$, if $\|v\| = 0$, then what must v be?

(iii) In $\mathbb{R}^4$, with the Minkowski metric, what is the sign of the vectors $(1,-1,1,0)$, $(2,0,0,1)$, $(0,0,1,1)$?

(iv) What are their magnitudes? Which vectors can be "normalized" to unit vectors by dividing by their lengths?

(v) A basis of mutually orthogonal unit vectors is called "orthonormal". Is the standard basis of $\mathbb{R}^4$ with the Minkowski metric orthonormal?

Index shifting

Suppose we have a vector space with an inner product G, with inverse G^{-1}.
Then we can introduce a streamlined notation for the related maps from V to V^* and back

$$\underbrace{V^\flat}_{\text{covector}}(u) \equiv G(u,V) = G_{ij}u^iV^j = \underbrace{(G_{ij}V^j)}u^i \qquad \text{flat for "down", lowering index}$$

$$[V^\flat]_i = G_{ij}V^j$$

$$f(\underbrace{g^\#}_{\text{vector}}) \equiv G^{-1}(f,g) = G^{ij}f_ig_j = f_i\underbrace{(G^{ij}g_j)} \qquad \text{Sharp for "up", raising the index.}$$

$$[g^\#]^i = G^{ij}g_j$$

Using the metric and its inverse we can associate a covector $V^\flat$ with each vector V and a vector $g^\#$ with each covector g. We have just introduced shorthand notation $\flat : V \to V^*$ and $\# : V^* \to V$ for the associated maps between the vector space and its dual. These two maps are inverses:

$$[(V^\flat)^\#]^i = G^{ij}(V^\flat)_j = G^{ij}G_{jk}V^k = \delta^i{}_kV^k = V^i$$

$$[(g^\#)^\flat]_i = G_{ij}(g^\#)^j = G_{ij}G^{jk}g_k = \delta^k{}_ig_k = g_i\,.$$

The innerproduct provides an "identification map" between a vector space and its dual. This turns out to be so useful that more shorthand notation is introduced:

$$V_i \equiv V^\flat(e_i) = G_{ij}V^j \qquad \text{"lowering the index"}$$

$$g^i \equiv \omega^i(g^\#) = G^{ij}g_j \qquad \text{"raising the index"}$$

We use the same letter for the corresponding covector or vector (called the kernel symbol, kernel in the sense that we add sub/superscripts to it) and just put the index in the right location. One then refers to the "contravariant" or "covariant" form of a vector, to distinguish the two.

EX. On $\mathbb{R}^n$ with the standard basis $\{e_i\}$ and the standard (dot) inner product $G_{ij} = \delta_{ij}$, $G^{ij} = \delta^{ij}$. The index shifting maps identify vectors and covectors with the same standard components

$$(V^\flat)_i \equiv V_i = \delta_{ij} V^j \qquad (\text{i.e., } V_i = V^i \text{ for each } i)$$
$$(f^\#)^i \equiv f^i = \delta^{ij} f_j \qquad (\text{i.e., } g^i = g_i \text{ for each } i).$$

Thus evaluation of a covector on a vector

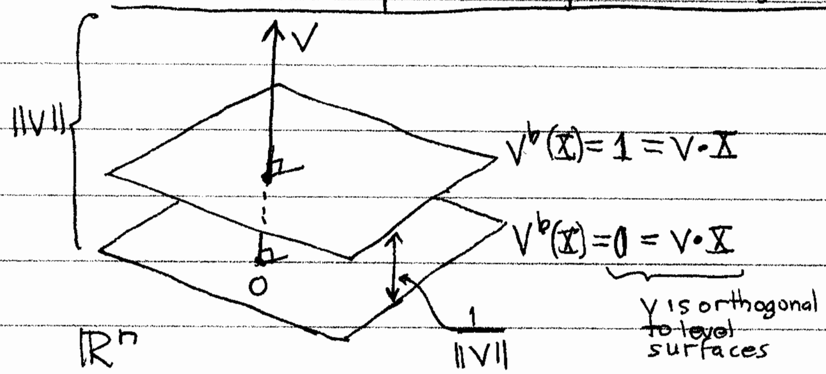
$$f(V) = f_i V^i = \delta_{ij} f^j V^i = V \cdot f^\#$$

is represented as the standard dot product of the vector with another vector whose components are the same as the covector.

In this way linearity is converted into geometry and one can ignore the distinction between the vector space and its dual and thus only use subscript indices. However, there is a catch. For everything to work, one has to use only orthonormal bases — otherwise things fall apart. (If the basis is not orthonormal, one no longer has the same components for a vector and its corresponding covector.) This turns out to be no problem for elementary linear algebra with its limited goals, but it is a problem if you want to go beyond that.

The relation $V^\flat(X) = V\cdot X = 0$ shows that the vector V is orthogonal to the level surfaces of the covector $V^\flat$.

Geometric interpretation of index shifting



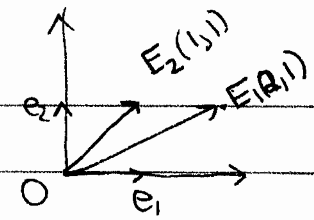
The relation
$$V^\flat(V) = V\cdot V = \|V\|^2$$
may be interpreted as stating that the vector V pierces $\|V\|^2$ "layers" (integer valued level surfaces) of the covector $V^\flat$.

Thus if V has length $\|V\|$, the "layer thickness" is $\frac{\|V\|}{\|V\|^2} = \frac{1}{\|V\|}$, i.e., the reciprocal of the length of the vector (i.e., a unit vector along V pierces $\|V\|$ layers). For a unit vector $\|V\| = 1$ and this separation is also 1.

In $\mathbb{R}^3$ we first learn to write an equation for a plane in the form
$$a(x-x_0) + b(y-y_0) + c(z-z_0) = 0$$
or $\quad \vec{N}\cdot(\vec{r}-\vec{r}_0) = 0 \quad$ where $\vec{N} = (a,b,c)$.

In fact (a,b,c) are the components of the covector, one of whose level surfaces is being described. This is then converted into a geometric statement about points whose difference vector from a reference point is perpendicular to ~~that plane~~ the vector whose components are the same as the coefficients of the linear function (covector).

PROBLEM $\mathbb{R}^2$



$\underline{A}^{-1} = (\underline{E_1}\ \underline{E_2})$

↑ columns = standard components of new basis

Let $e_{i'} \equiv E_i$ for the problem discussed above, to follow the change of basis ~~notation. Let $G = \delta_{ij}\omega^i \otimes \omega^j$ be the standard dot product tensor.~~

(i) Compute $G^{\circ}{}_{i'j'} = e_{i'} \cdot e_{j'}$ directly and from its transformation law.

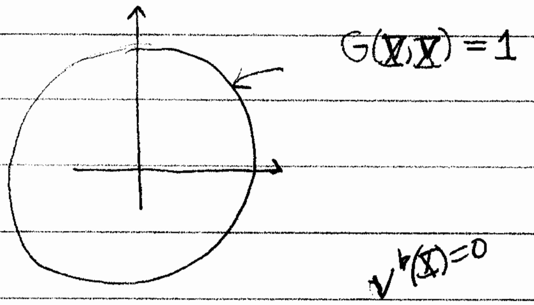
(ii) Compute the matrix $\underline{G}^{-1\prime} = (G^{i'j'})$ from its transformation law.
What is the matrix form of this transformation law?

(iii) The vector $Y = (0,2) = -2E_1 + 4E_2 = -2e_{1'} + 4e_{2'}$
has $Y^\flat = 2\omega^1$. Use $\underline{G}'$ to "lower" its indices in the new basis
verify that the expression for $Y^\flat$ in terms of $\omega^{1'}$ and $\omega^{2'}$ is $2\omega^1$.

Cute fact (Aside for your reading pleasure)

The relationship between a vector and covector determined by the Euclidean metric has a cute geometric interpretation.

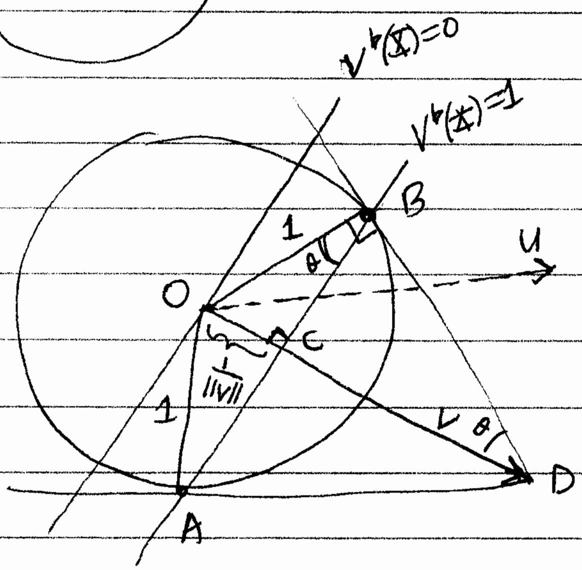
Consider the case of $\mathbb{R}^2$. The unit circle (all vectors of length 1) tells us everything we need to know about the Euclidean geometry of the metric tensor.



Suppose V is a vector with length bigger than 1 as in the diagram. Draw in the tangents AD and BD to the unit circle. Connect the points of tangency. Then OC has length $\frac{1}{\|V\|}$ since triangles ΔCOB and ΔCBD are similar so

$$\sin\theta = \frac{1}{\|V\|} = \frac{1/\|V\|}{1}$$

and $\overrightarrow{AB}$ is perpendicular to V.



Draw in a parallel line through the origin. Then these two parallel lines represent the covector $V^\flat = G(\,,V)$, since their separation is the reciprocal of the length of V, and they are orthogonal to V.

If we have another vector U, then the value of the metric on the pair $G(U,V) = V^\flat(U)$ is the number of "layers" of $V^\flat$ pierced by U, which is about 3 in the diagram. [This can be extended to $\|V\| < 1$ by inversion]

The same scheme works for any "positive definite" inner product on $\mathbb{R}^2$:

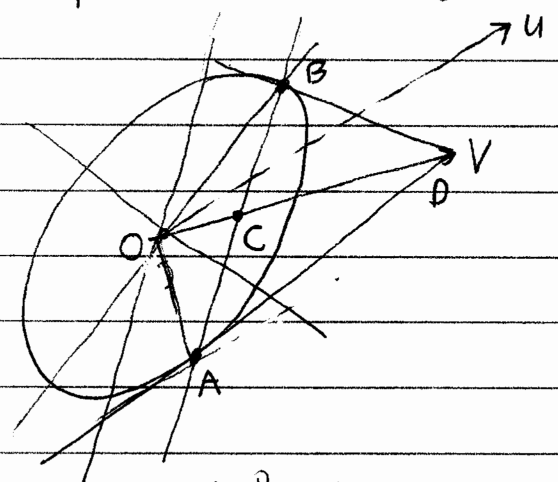
$$G = \underbrace{A}_{G_{11}}\,\omega^1\otimes\omega^1 + \underbrace{B}_{G_{12}=G_{21}}(\omega^1\otimes\omega^2+\omega^2\otimes\omega^1) + \underbrace{C}_{G_{22}}\,\omega^2\otimes\omega^2 \,, \quad \underbrace{A>0,\ C>0}_{\text{"positive definite"}}$$

[plus $AC-B^2>0$]

The unit circle for this metric, if $X=(x,y)$

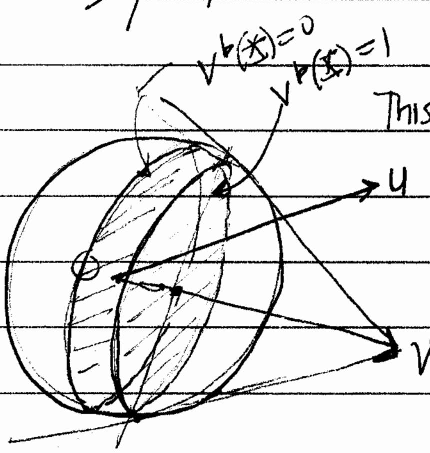
$$1 = G(X,X) = Ax^2+2Bxy+Cy^2$$

is an ellipse centered at the origin.



Exactly the same similar triangle geometry works here as well.

We can also interpret this as simply a representation of the Euclidean geometry of the plane but in nonorthogonal coordinates $\{x,y\}$ on the plane.



This also works with the unit "hypersphere" in $\mathbb{R}^n$ except one has a tangent "hypercone" with an $(n-2)$-sphere of tangency through which passes a hyperplane orthogonal to V. Together with the parallel hyperplane through the center of the hypersphere (the origin O), we get the representation of the covector $V^\flat$ and its value on another vector u in terms of the number of layers pierced.

Thus the unit hypersphere can "represent" an inner product, which is a symmetric "nondegenerate" (nonzero determinant of component matrix) $\binom{0}{2}$-tensor.
"Degenerate" (zero determinant) symmetric $\binom{0}{2}$-tensors have hypercylinder representations etc. We dont need these geometric interpretations, but sometimes they can be useful.

Index shifting conventions

In a situation where an inner product G is available and relevant to the kind of problem being described mathematically, we can extend the "index shifting" maps to any type of tensor.

A $\binom{p}{q}$ - tensor is said to have "rank" $(p+q)$ and have p contravariant indices (i.e. p covector arguments) and q covariant indices (i.e. q vector arguments).

For all tensors of a given total rank, we can establish a correspondence between tensors with different "index" positions. For example, if $p+q=2$, then we are dealing with $\binom{0}{2}$, $\binom{1}{1}$, or $\binom{2}{0}$ - tensors.

Suppose $T = T^i{}_j\, e_i \otimes \omega^j$. Then we can introduce two other tensors by

$$T^{ij} \equiv G^{jk} T^i{}_k \quad , \quad T_{ij} \equiv G_{ik} T^k{}_j \, .$$

These are related to each other in turn by

$$T^{ij} = G^{im} G^{jn} T_{mn} \, , \qquad T_{ij} = G_{im} G_{jn} T^{mn} \, .$$

For a given starting tensor T, we can interpret all three as different "representations" of the same physical object, but with different index arrangements. Of course this is a convenient fiction since a vector V and covector $V^\flat$ have completely different geometric interpretations, but those interpretations are related to each other in an interesting way.

We use the same kernel letter and let the index position distinguish between the different tensors of this family of related tensors.

For rank 3 tensors there are 8 different index positions:

$$\underbrace{T_{ijk}}_{\binom{0}{3}} ; \underbrace{T^i{}_{jk},\ T_i{}^j{}_k,\ T_{ij}{}^k}_{\binom{1}{2}} ; \underbrace{T_i{}^{jk},\ T^i{}_j{}^k,\ T^{ij}{}_k}_{\binom{2}{1}} ; \underbrace{T^{ijk}}_{\binom{3}{0}}$$

Now to distinguish between different index positions of different arguments of the tensor we have to suspend our convention of listing all the covector arguments first and the vector arguments last. This is not a problem since the index shifting turns out to be extremely useful.

Given any $\binom{p}{q}$-tensor there are two special members of the family of tensors related to it by index shifting, namely the "totally covariant" form of the tensor (all indices down) and the "totally contravariant" form of the tensor (all indices up)

$$
\begin{tikzcd}[row sep=small]
 & & T^{i\ldots j\ldots} \sim T^{\#} \\
T \sim T^{i\ldots}{}_{\underset{\sim}{j\ldots}} \arrow[urr] \arrow[drr] & & \\
 & & T_{i\ldots j\ldots} \sim T^{\flat}
\end{tikzcd}
$$

slide over ~~before raising them~~ or lowering the ~~upper indices~~

This extends the $\flat$ and $\#$ maps to arbitrary tensors, meaning respectively "raise all indices" and "lower all indices."

For the usual dot product on $\mathbb{R}^n$, using the standard basis, all of these tensors have the same numerical values for corresponding components, so one can always use the totally covariant form of a tensor accepting only vector arguments to discuss elementary linear algebra.

We can also introduce the magnitude and sign of a tensor just like that of a vector in terms of the totally covariant or contravariant form.

Define $\|T\| \geq 0$ and $\operatorname{sgn} T \in \{+,0,-\}$ by

$$
\begin{aligned}
(\operatorname{sgn} T)\,\|T\|^2 &= G^{im}G^{jn}\cdots T_{ij\ldots}T_{mn\ldots} = G_{im}G_{jn}\cdots T^{ij\ldots}T^{mn\ldots} \\
&= T_{ij\ldots}T^{ij\ldots} \quad [\text{with } \|T\| \equiv \operatorname{sgn} T \equiv 0 \text{ if this vanishes}]
\end{aligned}
$$

Since the space of $\binom{p}{q}$-tensors for fixed p and q is itself a vector space, it can have an inner product. This defines such an inner product induced by the inner product on the underlying vector space.

For tensors over $\mathbb{R}^n$ with the usual dot product, this inner product for tensors is very simple to describe. The sign is always positive, except for the zero tensor of a given valence (p and q !), and the magnitude is always positive (except for the zero tensor) and equal to the square root of the sum of the squares of all its components (just like for vectors!).

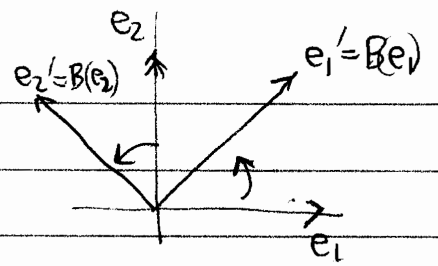
For example $\quad \|T\|^2 = \delta^{im}\delta^{jn} T_{ij} T_{mn} = T_{ij} T^{ij}$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} T_{ij} T^{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n} (T_{ij})^2 , \text{ since } T_{ij} = T^{ij} \text{ for}$$

~~each pair of index values (i, j).~~

Note that the inverse tensor $G^{-1} = G^{ij} e_i \otimes e_j$ defines an inner product on the dual space V^* thought of as a vector space in its own right, and this definition of the magnitude and sign of a covector is exactly the definition we introduced above for a vector space V, except now applied to the dual space.

PROBLEM $\mathbb{R}^2$



$\underline{B} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ is the matrix of an active rotation of the plane by $45°$ counterclockwise, so $\underline{A} = B^{-1} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ is the matrix by which vector indices transform.

The cartesian coordinates are the standard dual basis $x = \omega^1$, $y = \omega^2$, so

$$\omega^{i'} = A^{i'}{}_j \omega^j \quad \leftrightarrow \quad \begin{pmatrix} \omega^{1'} \\ \omega^{2'} \end{pmatrix} = A \begin{pmatrix} \omega^1 \\ \omega^2 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} \omega^1 \\ \omega^2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}}(\omega^1 + \omega^2) \\ \frac{1}{\sqrt{2}}(-\omega^1 + \omega^2) \end{pmatrix}$$

corresponds to the coordinate change $\begin{pmatrix} x' \\ y' \end{pmatrix} = A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}}(x+y) \\ \frac{1}{\sqrt{2}}(-x+y) \end{pmatrix}$.

Similarly $\underbrace{e_{i'}}_{\text{row matrix}} = \underbrace{e_j}_{\text{row matrix}} \underbrace{A^{-1j}{}_i}_{\text{adjacent} \to \text{matrix mult.}} \quad \leftrightarrow \quad (e_{1'}\ e_{2'}) = (e_1\ e_2)A^{-1} = (e_1\ e_2)\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$

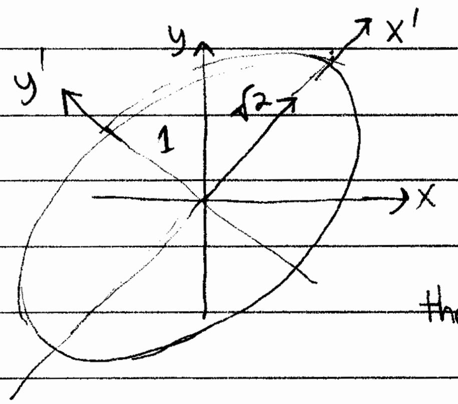
$$= \begin{pmatrix} \frac{1}{\sqrt{2}}(e_1 + e_2) \\ \frac{1}{\sqrt{2}}(-e_1 + e_2) \end{pmatrix}$$

Consider the symmetric tensor $H = H_{ij}\,\omega^i \otimes \omega^j$ with $\underline{H} = \begin{pmatrix} 3/2 & 1/2 \\ 1/2 & 3/2 \end{pmatrix}$.

(i) Verify that the change of basis leads to $\underline{H}' = \underline{A}^{-1T}\underline{H}\,\underline{A}^{-1} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ i.e., diagonalizes $\underline{H}$, while not changing the Euclidean inner product G: $\underline{G}' = \underline{A}^{-1T}\underline{I}\,\underline{A}^{-1} = \underline{I}$ (since $\{e_{i'}\}$ is also orthonormal).

(ii) compute the magnitude of H in each basis (magnitude wrt $\underline{G}$).

(iii) H may itself define an inner product. Why is its "unit circle" defined by the equation $1 = H(X, X) = \frac{3x^2 + 3y^2 + 2xy}{2} = 2(x')^2 + (y')^2$?



(iv) The semiaxes of this ellipse are 1 and $\sqrt{2}$. Notice that the ellipse lies inside the circle of radius equal to the magnitude $\sqrt{5}$ of H. Thus the magnitude does provide a useful bound on ~~the geometric representation of the inner product H.~~

partial evaluation of a tensor and index shifting

If we evaluate the inner product G only on its second argument "$G(\ ,v)$", then it still needs a vector in its first argument to produce a real number. This is a linear function of that argument, i.e., defines a covector, which is exactly $v^\flat$. We can write suggestively $v^\flat = G(\ ,v)$, for the partial evaluation of G on one argument. Similarly $f^\# = G(\ ,f)$.

We can partially evaluate any tensor on any number of arguments.

For example, if $\quad T = T_{ijk}\,\omega^i\otimes\omega^j\otimes\omega^k = \text{"}T(\ ,\ ,\ )\text{"}$

then $\quad T(\ ,v,\ ) \equiv T_{ijk}\,\omega^i\otimes\omega^k\,\omega^j(v) = T_{ijk}v^j\,\omega^i\otimes\omega^k$

makes sense as a way to represent partial evaluation on a single argument. Iteration of this extends it to any number of arguments.

contraction of tensors

For a $\binom{p}{q}$-tensor with at least one index of each type ($p\geq 1$, $q \geq 1$), one can select one upper index and one lower index and sum over them, reducing the number of free indices by 2 leading to a $\binom{p-1}{q-1}$-tensor. This is called contraction of the tensor on that pair of indices of opposite valence (one up, one down!)

For example, with a $\binom{1}{2}$-tensor $\quad T = T^i{}_{jk}\,e_i\otimes\omega^j\otimes\omega^k$

we get two covectors $\qquad T^k{}_{ki}\,\omega^i \qquad T^k{}_{ik}\,\omega^i \qquad$ from the two possible contractions of the single contravariant index with the two covariant indices.

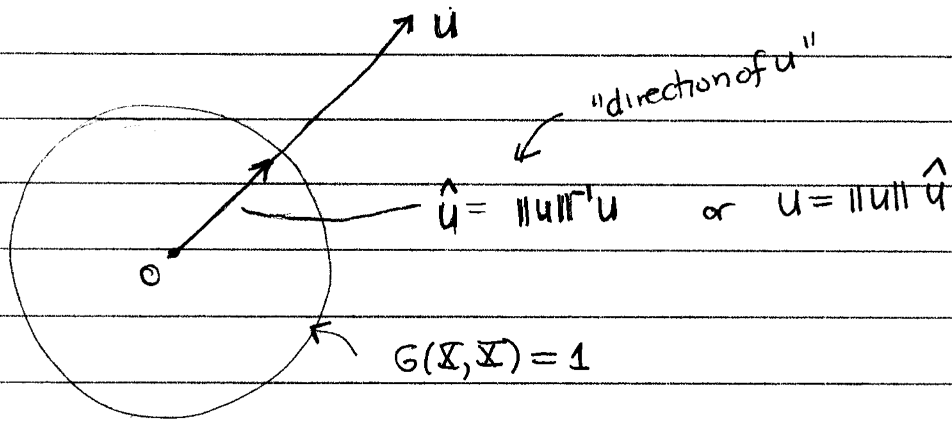
Remark for $\mathbb{R}^n$ with "positive-definite" inner product (sgn X always positive)

Using the symmetry property of our inner product $G(X,Y) = G(Y,X)$

$$G(X+Y, X+Y) = G(X,X) + G(Y,Y) + 2\,G(X,Y)$$

$$G(X,Y) = \frac{G(X+Y, X+Y) - G(X,X) - G(Y,Y)}{2}$$

we can determine all inner product values from self-inner product values, which explains how the unit sphere $G(X,X)=1$ can contain all the information about the inner product, even angle information.



Projection to unit sphere by normalization (for vectors whose sign is not zero).

$$G(X,Y) = \|X\|\,\|Y\|\; G(\hat{X},\hat{Y})$$

or $\quad X\cdot Y = \|X\|\,\|Y\|\; \hat{X}\cdot\hat{Y}$ $\qquad$ using the dot product notation.

For unit vectors the above relation becomes

$$\underbrace{G(\hat{X},\hat{Y})}_{\equiv \cos\theta} = \tfrac{1}{2}\left[\, G(\hat{X}+\hat{Y}, \hat{X}+\hat{Y}) - 1 - 1 \right] = \tfrac{1}{2}\underbrace{\|\hat{X}+\hat{Y}\|^2}_{\in[0,2]} - 1 \in [-1,1]$$

$\equiv \cos\theta$ defines cosine of angle between two directions

$\in [0,2]$ by Euclidean geometry

0 when $\hat{X} = -\hat{Y}$

2 when $\hat{X} = \hat{Y}$

For inner products with nonpositive sign values for vectors, this argument must be revised. More later.

Problem Let V be the n^2-dimensional vector space of $n\times n$ real matrices.
Define the standard basis by

$$\underline{A} = A^i{}_j e^j{}_i \quad \text{where} \quad e^j{}_i = n\times n \text{ matrix with 1 entry in } i\text{th row, } j\text{th column, zeros elsewhere}$$

row index, column index

Then

$$\begin{aligned}\underline{A} = {} & A^1{}_1 e^1{}_1 + A^1{}_2 e^2{}_1 + \cdots + A^1{}_n e^n{}_1 \equiv u^1 E_1 + u^2 E_2 + \cdots + u^n E_n \\ & + A^2{}_1 e^1{}_2 + \cdots \qquad + A^2{}_n e^n{}_2 \qquad + u^{n+1} E_{n+1} + \cdots + u^{2n} E_{2n} \\ & \vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots \\ & + A^n{}_1 e^1{}_n + \cdots + A^n{}_n e^n{}_n \qquad + u^{(n-1)n+1} E_{(n-1)n+1} + \cdots + u^{n^2} E_{n^2}\end{aligned}$$

defines an isomorphism $\underline{A} \in V \mapsto (u^1, \ldots, u^{n^2}) \in \mathbb{R}^{n^2}$
with $\mathbb{R}^{n^2}$, mapping this basis onto the standard basis of that space.
However, the original matrix notation is more useful because of matrix multiplication.

(i) If the dual basis is defined by $\omega^i{}_j(e^m{}_n) = \delta^i{}_n \delta^m{}_j$,
how are the components $A^i{}_j$ related to them?

(ii) Show that the matrix product law $e^i{}_j e^m{}_n = \delta^i{}_n e^m{}_j$
for the basis matrices extends by linearity to the usual index formula
for matrix multiplication $[\underline{A}\,\underline{B}]^i{}_j = A^i{}_k B^k{}_j$.

(iii) Define two inner products on V by

$$G(\underline{A},\underline{B}) = \mathrm{Tr}\, \underline{A}^T \underline{B} = \mathrm{Tr}\, \underline{A}\,\underline{B}^T \quad \text{where } \underline{A}^T \text{ is the transposed matrix and}$$

$$\mathscr{G}(\underline{A},\underline{B}) = \mathrm{Tr}\, \underline{A}\,\underline{B} = A^i{}_j B^j{}_i \quad \mathrm{Tr}\,\underline{A} \text{ is the trace of the matrix } A^i{}_i$$

$$\boxed{\text{Recall: } \mathrm{Tr}\,\underline{A} = \mathrm{Tr}\,\underline{A}^T \qquad (\underline{A}\,\underline{B})^T = \underline{B}^T \underline{A}^T \qquad \mathrm{Tr}\,\underline{A}\,\underline{B} = \mathrm{Tr}\,\underline{B}\,\underline{A}}$$

If we write $[\underline{A}^T]^i{}_j = \delta_{jn} A^n{}_m \delta^{mi}$ in order to respect our index
conventions, then $G(\underline{A},\underline{B}) = [\underline{A}^T]^i{}_j B^j{}_i = \delta_{jn}\delta^{mi} A^n{}_m B^j{}_i$, and
$G(\underline{A},\underline{A}) = \delta_{jn}\delta^{mi} A^n{}_m A^j{}_i = \sum_{j=1}^{n}\sum_{i=1}^{n} (A^j{}_i)^2$ = sum of squares
of all entries of matrix. Thus G corresponds to the usual dot product on
$\mathbb{R}^{n^2}$ under the above correspondence. Make sure you understand this.

Note that $\mathscr{G}(A,B) = G(A^T, B)$.

TOPS FORM 7525

Litho in U.S.A.

(iv) Suppose $\underline{A} = \underline{A}^T$ is symmetric and $\underline{B} = -\underline{B}^T$ is antisymmetric. Using the Euclidean property of positive-definiteness $G(A,A) \geq 0$ with $G(A,A) = 0$ iff $A = 0$, then

$$\mathcal{G}(A,A) = G(A^T,A) = G(A,A) \geq 0$$
$$\mathcal{G}(B,B) = G(B^T,B) = -G(B,B) \leq 0$$

shows that $\operatorname{sgn} A = 1$, $\operatorname{sgn} B = -1$ for all nonzero symmetric and antisymmetric matrices respectively. Use a similar argument to show that $\underline{A}$ and $\underline{B}$ are orthogonal with respect to both inner products.

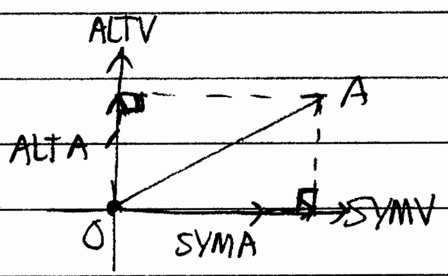
(v) Is the basis $\{\underline{e}^i{}_j\}$ of V orthonormal with respect to both inner products? Why?

(vi) The subspaces SYM V and ALT V of symmetric and antisymmetric matrices of V are each vector subspaces (why?), and every matrix can be written uniquely in terms of its symmetric and antisymmetric parts

$$\underline{A} = \underbrace{\mathrm{SYM}\,\underline{A}}_{=\frac{1}{2}(\underline{A}+\underline{A}^T)} + \underbrace{\mathrm{ALT}\,\underline{A}}_{=\frac{1}{2}(\underline{A}-\underline{A}^T)}$$

V is said to be a "direct sum" of these two vector subspaces. Their dimensions are

$$\left.\begin{aligned} \dim(\mathrm{SYM}\,V) &= \sum_{i=1}^{n} i = \frac{n(n+1)}{2} \\ \dim(\mathrm{ALT}\,V) &= \Big(\sum_{i=1}^{n} i\Big) - n = \frac{n(n-1)}{2} \end{aligned}\right\} \text{ why?}$$



The maps $A \mapsto \mathrm{SYM}\,A$, $A \mapsto \mathrm{ALT}\,A$ are projection maps associated with this direct sum. They are orthogonal with respect to both inner products, in the sense that they project onto orthogonal subspaces.

[Projection maps satisfy $P^2 = P$, $Q^2 = Q$, $PQ = QP = 0$ for a pair (P,Q) which projects onto two subspaces in a direct sum.]

TOPS FORM
7525

Litho In U.S.A.

(vii) Make the following definitions

$$\bar{E}^i{}_j = 2^{-1/2}(e^i{}_j - e^j{}_i) \qquad \bar{A}^i{}_j = 2^{-1/2}(A^i{}_j - A^j{}_i)$$

$$\hat{E}^i{}_j = \begin{cases} e^i{}_j & , i=j \\ 2^{-1/2}(e^i{}_j - e^j{}_i), & i\neq j \end{cases} \qquad \hat{A}^i{}_j = \begin{cases} A^i{}_j & , i=j \\ 2^{-1/2}(A^i{}_j - A^j{}_i), & i\neq j \end{cases}$$

Then $\underline{A} = A^i{}_j \underline{e}^j{}_i = \sum_{i\leq j} \hat{A}^i{}_j \hat{\underline{E}}^j{}_i + \sum_{i<j} \bar{A}^i{}_j \bar{\underline{E}}^j{}_i$

shows that $\{\hat{\underline{E}}^j{}_i\}_{i\leq j} \cup \{\bar{\underline{E}}^j{}_i\}_{i<j}$ is a basis of V adapted to the "orthogonal" direct sum into symmetric and antisymmetric matrices.

Evaluate both inner products of the pairs $(\hat{E}^i{}_j, \hat{E}^m{}_n)$, $(\hat{E}^i{}_j, \bar{E}^m{}_n)$, $(\bar{E}^i{}_j, \bar{E}^m{}_n)$.

What are the lengths of these basis vectors?

What are their signs with respect to each inner product?

What kind of basis is this with respect to either inner product?

(viii) If we introduce the covector index positioning by

$f = f^i{}_j\, \omega^j{}_i$, $f^i{}_j = f(\underline{e}^i{}_j)$, $\omega^i{}_j(e^m{}_n) = \delta^i{}_n \delta^m{}_j$ (duality),

then we can associate a vector $\underline{F} = f^i{}_j \underline{e}^j{}_i$ with each such covector.

Show that $f(\underline{A}) = \mathrm{Tr}\, \underline{F}\underline{A} = \mathcal{G}(\underline{F}, \underline{A})$,

i.e., $\underline{F} = f^\#$ with respect to $\mathcal{G}$.

[Remark: If we had instead used the notation

$f = f_i{}^j \omega^i{}_j$, $\omega^i{}_j(e^m{}_n) = \delta^i{}_n \delta^m{}_j$, $f_i{}^j = f(\underline{e}^j{}_i)$,

we would have found instead $f(\underline{A}) = \mathrm{Tr}(\underline{F}^T \underline{A}) = G(\underline{F}, \underline{A})$

if we let $\underline{F} = f_i{}^j e^i{}_j$.

We could have also used the alternate notation $\underline{A} = A^i{}_j e_i{}^j$ from the beginning, which would have resulted in further changes. It is important to realize that a choice of notation implies certain implicit choices not obvious at first. Even other choices $A = A_{ij} e^{ij}$ or $A = A^{ij} e_{ij}$ are possible.]

(ix) Suppose we define $H = H^i{}_j{}^m{}_n\, \omega^j{}_i \otimes \omega^n{}_m$, $H^i{}_j{}^m{}_n = H(e^j{}_i, e^n{}_m)$ for any $\binom{0}{2}$-tensor over V. What are the components of G and $\mathscr{S}$ using this notation?

(x) $\mathscr{S} = \mathrm{Tr}\,\underline{A}\,\underline{B}$ is a $\binom{0}{2}$-tensor. Why?

For the same reason, for each positive integer p, the following defines a $\binom{0}{p}$-tensor over V: $T^{(p)}(\underbrace{\underline{A}, \underline{B}, \ldots, \underline{C}}_{p \text{ vector arguments}}) = \mathrm{Tr}(\underbrace{\underline{A}\,\underline{B} \cdots \underline{C}}_{p \text{ factors}})$.

$T^{(1)}$ is a covector. Express it in terms of the dual basis.

Note that the cyclic property of the trace $\mathrm{Tr}\, AB \ldots CD = \mathrm{Tr}\, B \ldots CDA = \cdots$ *implies certain symmetries of these tensors. It makes* $T^{(2)} = \mathscr{S}$ *symmetric.*

(xi) If we define $D^{(p)}(\underline{A}, \underline{B}, \ldots, \underline{C}) = \det(\underline{A}\,\underline{B} \cdots \underline{C})$, is this a tensor? Why?

(xii) *Sketchy remark for your mathematical interest (just read for pleasure)*

The inner product $g(\underline{A},\underline{B}) = \mathrm{Tr}\,\underline{A}\underline{B} - \mathrm{Tr}\,\underline{A}\,\mathrm{Tr}\,\underline{B}$ only differs from $\mathscr{G} = \mathrm{Tr}\,\underline{A}\underline{B}$ on the symmetric matrices since antisymmetric matrices have zero trace. (Why?) The symmetric matrices themselves may be decomposed into an offdiagonal subspace (again zero trace) and a diagonal subspace, while the diagonal subspace itself can be decomposed into the tracefree subspace and the 1-dimensional "pure trace" subspace of multiples of the identity matrix

$$\underline{A} = \underbrace{\left(\tfrac{1}{n}\mathrm{Tr}\,\underline{A}\right)\underline{I}} + \underbrace{\left[\underline{A} - \left(\tfrac{1}{n}\mathrm{Tr}\,\underline{A}\right)\underline{I}\right]} = \underbrace{\sum_{i=j} A^i{}_j \underline{e}^j{}_i} + \underbrace{\sum_{i\neq j} A^i{}_j \underline{e}^j{}_i}$$

$$\overset{(1)}{=} \underline{A}^{\text{trace}} + \underline{A}^{\text{tracefree, sym}} \overset{(2)}{=} \underline{A}^{\text{diagonal}} + \underline{A}^{\text{offdiagonal, sym}}$$

$$\overset{(3)}{=} \underbrace{\underline{A}^{\text{trace}} + \underline{A}^{\text{tracefree, diagonal}}}_{} + \underline{A}^{\text{offdiagonal, sym}}$$

$$= \underline{A}^{\text{diagonal}} + \underline{A}^{\text{offdiagonal, sym}}$$

Each of these three decompositions (1), (2), (3) are orthogonal decompositions of the subspace of symmetric matrices with respect to $\mathscr{G}$ (which coincides with G for symmetric matrices — they differ in sign on the antisymmetric matrices), while the symmetric & antisymmetric matrices are orthogonal with respect to both $\mathscr{G}$ and G so it extends to an orthogonal decomposition of V itself. Anyway the new inner product g only differs from G and $\mathscr{G}$ on the 1-dimensional space of pure trace matrices, which has negative sign with respect to g (G and $\mathscr{G}$ have all positive signs for symmetric matrices.)

$$\{\underline{I}\} \cup \{\underline{e}^i{}_i - \tfrac{1}{n}\underline{I}\}_{i=1,\dots,n-1}$$ is an orthogonal basis of the diagonal subspace adapted to this pure trace/tracefree decomposition, which has only one basis vector with a negative sign. Such inner products where the orthonormal bases have only one negative sign are called LORENTZIAN. (like 4-dimensional Minkowski spacetime)

TOPS FORM 7525

Litho In U.S.A.

Without pursuing the details, you can see that just pushing on some simple familiar properties of matrices leads to an extremely rich structure complete with geometry. In fact the space of symmetric matrices with nonzero determinant is an open subspace of the set of all symmetric matrices and may be interpreted as a "curved space" of all possible (symmetric) innerproducts on $\mathbb{R}^n$. This turns out to play a key role in the structure of the complicated nonlinear coupled partial differential equations of general relativity called Einstein's equations.

If you really like mathematics, you can see that by properly recognizing mathematical structure and adapting notation to it, one can create out of nothing a beautiful arena of geometry— which in fact is not just idle games playing but often has important applications in physical science. On the other hand, sweeping the structure under the rug in order to arrive immediately at calculational algorithms (as unfortunately we must in a one semester linear algebra course) completely hides this structure and the "beauty". Our goal is simply to begin to appreciate how this can be uncovered and see how it applies to the geometry of "curved spaces", which itself has enormous importance in the physical sciences.