5

## WHAT YOU GET FROM A TYPICAL LINEAR ALGEBRA COURSE: NOTATION

$\mathbb{R}^n = \{ (x_1, \dots, x_n) \mid x_\alpha \in \mathbb{R} \}$ $\qquad \vec{x} = (x_1, \dots, x_n) = \sum_{\alpha=1}^{n} x_\alpha e_\alpha$

standard basis: $\qquad e_\alpha = (0, \dots, \underset{\substack{\uparrow \\ \alpha\text{th slot}}}{1}, \dots, 0) \quad , \quad 1 \le \alpha \le n$

$\{x_\alpha\}_{\alpha=1,\dots,n}$ are the components of the vector $\vec{x}$ wrt the standard basis.

Suppose you have a linear map $A : \mathbb{R}^n \to \mathbb{R}^m$.

Let $i, j, k, \dots$ denote indices assuming the values $1, \dots, m$.

Let $\{e_{i'}\}_{i=1,\dots,m}$ denote the standard basis of $\mathbb{R}^m$ to avoid confusion.

The matrix of $A$ is defined by:

$$A e_\alpha = \sum_{i=1}^{m} A_{i\alpha} e_{i'}$$

so that
$$\begin{aligned} A\vec{x} = A\Big(\sum_{\alpha=1}^{n} x_\alpha e_\alpha\Big) &= \sum_{\alpha=1}^{n} x_\alpha A e_\alpha \qquad \text{by linearity} \\ &= \sum_{\alpha=1}^{n} x_\alpha \sum_{i=1}^{n} A_{i\alpha} e_{i'} \\ &= \sum_{i=1}^{m} \underbrace{\Big( \sum_{\alpha=1}^{n} A_{i\alpha} x_\alpha \Big)}_{\equiv x_{i'}} e_{i'} \equiv \vec{x}\,' \end{aligned}$$

hence $\qquad \underline{x}' = \underline{A}\,\underline{x} \qquad$ in matrix notation,

if we agree that $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ is the column matrix associated with $\vec{x} = (x_1, \dots, x_n)$ and similarly for $\vec{x}\,'$, and

$$\underline{A} = (A_{i\alpha}) \qquad \begin{array}{l} i = 1, \dots, m \text{ (rows)} \\ \alpha = 1, \dots, n \text{ (columns)} \end{array}$$

is the matrix associated with the array of numbers $\{A_{i\alpha}\}_{\substack{i=1,\dots,m \\ \alpha=1,\dots,n}}$, and use matrix multiplication to multiply $\underline{A}$ and $\underline{\vec{x}}$.

If one has the composition of two linear maps, one matrix multiplies their matrices to get the matrix of the composition, and so on.

If $A: \mathbb{R}^n \to \mathbb{R}^n$, then one has a square matrix and

$$A e_\alpha = \sum_{\beta=1}^{n} A_{\beta\alpha} e_\beta$$

$$x_{\alpha'} = \sum_{\beta=1}^{n} A_{\alpha\beta} x_\beta .$$

Define the determinant of a square matrix by

$$\det \underline{A} = \sum_{\text{permutations } \pi \text{ of } (1,\dots,n)} (\operatorname{sgn}\pi) A_{1\,\pi(1)} \cdots A_{n\,\pi(n)}$$

or $$\det \underline{A} = \sum_{\text{permutations } \pi \text{ of } (1,\dots,n)} (\operatorname{sgn}\pi) A_{\pi(1)\,1} \cdots A_{\pi(n)\,1}$$

which is a consequence of the fact $\det \underline{A}^T = \det \underline{A}$, where the transpose matrix is $(A^T)_{\alpha\beta} = A_{\beta\alpha}$.

- A linear transformation $A: \mathbb{R}^n \to \mathbb{R}^n$ or its matrix $\underline{A}$ are called singular/nonsingular or degenerate/nondegenerate or noninvertible/invertible if $\det \underline{A}$ is zero/nonzero.

  If one has the composition of two such linear transformations $A \circ B$, then $\underline{AB}$ is the matrix of $A \circ B$ and their determinant factors

$$\det \underline{AB} = \det \underline{A} \, \det \underline{B} .$$

- For a nonsingular matrix $\underline{A}$, a unique left/right inverse matrix $\underline{A}^{-1}$ (and inverse linear transformation) exists for which

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

where $\underline{1} = (\delta_{\alpha\beta})$ is the unit matrix and

$$\delta_{\alpha\beta} = \begin{cases} 1 & \alpha = \beta \\ 0 & \alpha \neq \beta \end{cases}$$ is the Kronecker delta symbol.

This immediately implies $\det \underline{A}^{-1} = (\det \underline{A})^{-1}$.

SHORTHAND. Since typically many summations will be involved in what we do, we adopt the Einstein summation convention:

1) If more than 2 indices are repeated in an expression, the summation convention is suspended and one must use the $\Sigma$ notation if a sum is desired.

2) If 2 indices are repeated in an expression, a sum is implied over their allowed range of values. Thus we need to adopt certain letters for certain allowed ranges in order to code the allowed range into our notation.
A repeated index is called a "dummy index" and may be replaced at will with affecting the expression.

3) If an index is not repeated in an expression, the expression is ~~assum~~ assumed to stand for all expressions for which the index assumes an allowed value.

Thus
$$\vec{X} = x_\alpha e_\alpha \ , \qquad \vec{X}' = x_{i'} e_{i'} \quad \text{or} \quad \vec{X}' = x_{\alpha}' e_{\alpha}'$$
$$A e_\alpha = A_{i\alpha} e_i \quad \text{or} \quad A e_\alpha = A_{\gamma\alpha} e_\beta$$
$$x'_{i'} = A_{i\alpha} x_\alpha \quad \text{or} \quad x_{\alpha'} = A_{\alpha\beta} x_\beta$$
making all these expressions more compact.
$$A^{-1}{}_{\alpha\gamma} A_{\gamma\beta} = \delta_{\alpha\beta} = A_{\alpha\gamma} A^{-1}{}_{\gamma\beta}$$

To deal with determinants in the same way we define.

$$\epsilon_{\alpha_1 \cdots \alpha_n} = \begin{cases} \underbrace{\operatorname{sgn}\begin{pmatrix} 1 & \cdots & n \\ \alpha_1 & \cdots & \alpha_n \end{pmatrix}}_{\substack{\text{obvious short hand} \\ \text{for permutation } \pi \text{ taking } (1,\ldots,n) \text{ to } (\alpha_1,\ldots,\alpha_n)}} & \text{if } (\alpha_1,\ldots,\alpha_n) = \pi(1,\ldots,n) \\ 0 & \text{if } (\alpha_1,\ldots,\alpha_n) \text{ is not a permutation of } (1,\ldots,n) \end{cases}$$

so with the summation convention the determinant also becomes compact:
$$\det \underline{A} = \epsilon_{\alpha_1 \ldots \alpha_n} A_{1\alpha_1} \cdots A_{n\alpha_n} = \epsilon_{\alpha_1 \ldots \alpha_n} A_{\alpha_1 1} \cdots A_{\alpha_n n}.$$

$\mathbb{R}^n$ has a standard Euclidean inner product (the "dot product")

$$\vec{x}\cdot\vec{y} = x_\alpha y_\alpha = \vec{y}\cdot\vec{x}$$

The standard basis vectors satisfy

$$e_\alpha \cdot e_\beta = \delta_{\alpha\beta} ,$$

so $\vec{x}\cdot\vec{y} = (x_\alpha e_\alpha)\cdot(y_\beta e_\beta) = x_\alpha y_\beta \delta_{\alpha\beta} = x_\alpha y_\alpha = x_\beta y_\beta .$

The affect of summing an index of the Kronecker delta against another index is to replace that other index with the other index of the Kronecker delta.

This may be used to interpret matrix multiplication:

$$\underset{m\times p}{A_{\alpha\beta}} = \underset{m\times n}{B_{\alpha\gamma}} \underset{n\times p}{C_{\gamma\beta}}$$

The $\alpha\beta$' entry of the product matrix is just the inner product of the $\alpha$th row vector of the first factor with the $\beta$th column vector of the second, where each such row and column may be considered a point of $\mathbb{R}^n$ in a natural way.

The quantity $\|\vec{x}\| = (\vec{x}\cdot\vec{x})^{1/2}$ is called the length of the vector. Any nonzero vector can be "normalized" to a unit vector (unit length) by dividing by its length

$$\hat{x} = \frac{\vec{x}}{\|\vec{x}\|} , \quad \hat{x}\cdot\hat{x} = 1 .$$

The relative angle between two nonzero vectors only depends on their corresponding unit vectors

$$\cos\theta = \hat{x}\cdot\hat{y} \quad \text{or} \quad \vec{x}\cdot\vec{y} = \|\vec{x}\|\,\|\vec{y}\| \cos\theta .$$

Whenever any two vectors have zero inner product, they are called orthogonal. When they are both nonzero, this means they determine perpendicular directions, i.e., their relative angle is $\frac{\pi}{2}$ radians or $90°$.

The condition $e_\alpha \cdot e_\beta = \delta_{\alpha\beta}$ makes the standard basis an orthonormal basis with respect to the dot product, ie. a basis of mutually orthogonal unit vectors.

The unit vectors parametrize the space of directions.
The scalar component of any vector $\vec{X}$ along the direction $\hat{u}$ is just $\vec{X}\cdot\hat{u} = \|\vec{X}\|\cos\theta$,
while the vector component is $(\vec{X}\cdot\hat{u})\,\hat{u}$, also called the projection of $\vec{X}$ along $\hat{u}$.

Unfortunately standard elementary linear algebra obscures the ~~full~~ rich structure of linearity and mixes it up with the Euclidean geometry of $\mathbb{R}^n$ by avoiding the introduction of the idea of dual spaces and their consequences. This confusion of the distinct concepts of an inner product between vectors and a natural scalar product ~~between~~ between a vector and a linear function (just evaluation) is built into the index notation which keeps all indices at the subscript level rather than distinguishing between subscript and superscript positioning. To separate these two different ideas, one must refine our index conventions to distinguish between "upper" and "lower" indices.

SUMMATION CONVENTION: A repeated pair of indices in an expression must consist of an upper and a lower index.

If we decide $\{e_\alpha\}$ will represent the standard basis of $\mathbb{R}^n$, then the components of vectors with respect to this basis must have an upper index

$$\vec{X} = X^\alpha e_\alpha \,.$$

Now when we make the dot product of two vectors

$$\vec{X}\cdot\vec{Y} = (X^\alpha e_\alpha)\cdot(Y^\beta e_\beta) = X^\alpha Y^\beta e_\alpha \cdot e_\beta = X^\alpha Y^\beta \delta_{\alpha\beta} = \delta_{\alpha\beta} X^\alpha Y^\beta$$

we can no longer eliminate the Kronecker delta since both remaining summed indices would be up. The mandatory presence of the Kronecker delta with two lower indices therefore signals the presence of a dot product. ~~Furthermore~~ However, we could make the definition $X_\beta = \delta_{\alpha\beta} X^\alpha = \delta_{\beta\alpha} X^\alpha$ to absorb the Kronecker delta into the notation without breaking the new index convention

$$\vec{X}\cdot\vec{Y} = \delta_{\alpha\beta} X^\alpha Y^\beta = X_\beta Y^\beta ,$$

but now we must consider $X_\beta$ distinct from $X^\beta$. What does this mean? And what is the nature of the product $X_\beta Y^\beta$?

Such a set is linearly independent if $C^A X_{(A)} = 0$ implies that every coefficient $C^A = 0$; otherwise it is called linearly dependent. A set of n linearly independent vectors is called a basis and any vector may be expressed uniquely as a linear combination of these vectors, the expansion coefficients called the components of it with respect to the basis.

Let us re-examine linearity in our new notation.

[The arrow notation is dropped.]

Suppose $\{X_{(A)}\}_{A=1,\dots,p}$ is a set of vectors in $\mathbb{R}^n$. The parenthesis subscript notation reminds us that the index just is a label of the set rather than a component of something. If $C^A$ are any p real numbers then $C^A X_{(A)}$ is a linear combination of this set of vectors.

---

Suppose $f : \mathbb{R}^n \to \mathbb{R}^m$ is a linear map. Linearity just means that evaluating f commutes with the operation of taking a linear combination

$$f(C^A X_{(A)}) = C^A f(X_{(A)}).$$

Linearity completely determines a linear map from its values on the standard basis vectors

$$f(X) = f(X^\alpha e_\alpha) = X^\alpha \underbrace{f(e_\alpha)}_{\equiv f_\alpha} = f_\alpha X^\alpha .$$

The quantities $f_\alpha \in \mathbb{R}^m$ are just the images of the standard basis vectors. The index summation which occurs on the right hand side of this equation is equivalent to evaluating the linear function in terms of components taken with respect to the basis $\{e_\alpha\}$.

Suppose we denote the standard basis of $\mathbb{R}^m$ by $\{e'_i\}$ to distinguish it from $\{e_\alpha\}$ on $\mathbb{R}^n$, where $i, j, k, \dots$ assume the range of values $1, \dots, m$. The image vectors $f_\alpha$ can be expressed in terms of the basis

$$f_\alpha = f^i{}_\alpha e'_i$$

$$f(X) = (f^i{}_\alpha X^\alpha) e'_i = Y^i e'_i$$

$$Y^i = f^i{}_\alpha X^\alpha \quad \text{or} \quad \underline{Y} = \underline{f}\,\underline{X}$$

if we let $\underline{X}$ and $\underline{Y}$ stand for the column matrices ~~correspo~~ of components of X and Y, and $\underline{f}$ ~~be the~~ $= (f^i{}_\alpha)$ the matrix whose rows are labeled by the first (upper) index and columns by the second (lower) index.

If $m=n$, then the prime is no longer necessary and all indices revert to Greek letters $\quad Y^\beta = f^\beta{}_\alpha X^\alpha$.

The identity transformation with identity matrix $\underline{1}$ must therefore be written $\quad \underline{1} = (\delta^\alpha{}_\beta)$, where now the Kronecker delta indices are in the position for a linear map.

Notice that the ~~column~~ $\alpha$th column of $\underline{f} = (f^i{}_\alpha)$ correspond to the components of the image vector of $e_\alpha$ in the new basis $\{e'_i\}$, or when $m=n$ and $\underline{f} = (f^\beta{}_\alpha)$, the components of the image of $e_\alpha$ in the same (standard) basis.

Consider the case $m=1$ of real valued linear functions on $\mathbb{R}^n$. Suppose $\sigma$ is such a function. Then

$$\sigma(X) = \sigma_\alpha X^\alpha, \qquad \sigma_\alpha \equiv \sigma(e_\alpha).$$

The space of such functions is itself a vector space since a linear combination of linear functions is again a linear function (when defined in the obvious way) and all the necessary rules for vector addition and scalar multiplication hold. This vector space is called the dual space and is denoted by $\mathbb{R}^{n*}$. It too has a standard basis, denoted by $\{\omega^\alpha\}$, called the basis dual to $\{e_\alpha\}$ or just the dual basis.

~~Make the definition~~

Define the $n$ linear functions $\omega^\alpha$ by

$$\omega^\alpha(e_\beta) = \delta^\alpha{}_\beta. \qquad \text{(duality relations)}$$

The linear function $\omega^\alpha$ for a fixed value of $\alpha$ gives zero on all basis vectors except the one with the same index, which gives a value of 1. The values on an arbitrary vector follow from linearity

$$\omega^\alpha(X) = \omega^\alpha(X^\beta e_\beta) = X^\beta \omega^\alpha(e_\beta) = \delta^\alpha{}_\beta X^\beta = X^\alpha.$$

The $\alpha$th dual basis vector picks out the $\alpha$th component of $X$.

But if $\sigma$ is an arbitrary such linear function

$$\sigma(X) = \sigma_\alpha X^\alpha = \sigma_\alpha \omega^\alpha(X) = (\sigma_\alpha \omega^\alpha)(X)$$

and since any two linear functions which have the same value on every vector must coincide

$$\sigma = \sigma_\alpha \omega^\alpha$$

which shows that $\{\omega^\alpha\}$ is a spanning set for the dual space, i.e., any element of the space can be expressed as a linear combination of the set. To prove that they form a basis, one must show linear independence, but if

$$C_\alpha \omega^\alpha = 0 \quad \text{then } 0 = C_\alpha \omega^\alpha(e_\beta) = C_\alpha \delta^\alpha{}_\beta = C_\beta$$

shows that each coefficient must vanish, establishing linear independence. Thus the dual space $\mathbb{R}^{n*}$ is also n-dimensional.

Thus any element of the dual space, called a covector, has the expansion

$$\sigma = \sigma_\alpha \omega^\alpha \qquad \sigma_\alpha = \sigma(e_\alpha)$$

where the components with respect to the dual basis, hereafter simply called the components with respect to the basis $\{e_\alpha\}$, are just the values of the covector on the basis vectors. Similary

$$X = X^\alpha e_\alpha \,, \qquad X^\alpha = \omega^\alpha(X)$$

shows that the components of a vector may be obtained by evaluating the dual basis covectors on the vector. Notice how the index positioning is now telling us the origin of an index.

If $f : \mathbb{R}^n \to \mathbb{R}^m$ is a linear map, and we introduce the dual basis $\{\omega'^i\}$ for $\mathbb{R}^m$, then

$$Y^i = f^i{}_\alpha X^\alpha \quad \text{where} \quad f(e_\alpha) = f^i{}_\alpha e'_i$$

means $\quad f^i{}_\alpha = \omega'^i(f(e_\alpha))$.

Again the index positioning reveals the origin and different function of the two indices. The second (lower) index is associated with the linearity of the map and its corresponding summation ~~just~~ is equivalent to evaluation of the map in terms of components with respect to the basis of the domain space $\mathbb{R}^n$. The upper (first) index can be used even for a nonlinear map.

When $m=n$ and $\omega'^i$ is just $\omega^\alpha$, one has

$$f^\beta{}_\alpha = \omega^\beta(f(e_\alpha)).$$

Now what about the dot product? $X\cdot Y = \delta_{\alpha\beta} X^\alpha Y^\beta$
Suppose we fix $X$, ~~then~~ we obtain a real valued linear function on $\mathbb{R}^n$, i.e., a covector "$X\cdot$" $= \delta_{\alpha\beta} X^\beta \omega^\alpha$ whose value on $Y$ is just $X\cdot Y$, and whose components are $\delta_{\alpha\beta}X^\beta \equiv X_\alpha$, ie., numerically the same as the components of the original vector. This is an obvious isomorphism between $\mathbb{R}^n$ and its dual space, but not a natural one because it involves a particular choice of basis. [Something is called "natural" if no arbitrary choice is involved.] As we will see later, such a correspondence in a different basis will not agree in general with this correspondence. [The arbitrary choice here involves the standard Euclidean inner product if one thinks of $X \Leftrightarrow$ "$X\cdot$" or the standard basis if one thinks $X_1 = X^1, \dots, X_n = X^n$.]

Suppose we explore a bit changing the basis from the standard one. This may be accomplished using a linear transformation $B$ of $\mathbb{R}^n$ which is nonsingular, ie., $\det \underline{B} \neq 0$. In this case the image vectors $e'_\alpha \equiv B(e_\alpha) = B^\beta{}_\alpha e_\beta$ are ~~nonsingular and~~ linearly independent and therefore form a new basis for $\mathbb{R}^n$. Define the new dual basis $\{\omega'^\alpha\}$ by the same duality relations as before

$$\delta^\alpha{}_\beta = \omega'^\alpha(e'_\beta) = \omega'^\alpha(B^\gamma{}_\beta e_\gamma)$$
$$= \omega'^\alpha(e_\gamma)\, B^\gamma{}_\beta$$

which means that the matrix $\omega'^\alpha(e_\gamma) = A^\alpha{}_\gamma$ of components of the new dual basis must be the inverse matrix $\underline{A} = \underline{B}^{-1}$.

Any vector may be expressed in either basis

$$X = X^\alpha e_\alpha = X'^\beta e'_\beta$$

$$X'^\alpha = \omega'^\alpha(\underline{X}) = A^\alpha{}_\beta\,\omega^\beta(\underline{X}) = A^\alpha{}_\beta \underline{X}^\beta \qquad \text{or} \quad \underline{X}' = \underline{A}\,\underline{X}$$

which can be inverted by matrix multiplication by the inverse

$$X^\alpha = A^{-1\alpha}{}_\beta\, X'^\beta \qquad \underline{X} = \underline{A}^{-1}\underline{X}'.$$

Similarly for a covector

$$\sigma = \sigma_\alpha \omega^\alpha = \sigma'_\alpha \omega'^\alpha$$

$$\sigma_{\alpha'} = \sigma(e_{\alpha'}) = \sigma(B^\beta{}_\alpha e_\beta) = \sigma_\beta B^\beta{}_\alpha = \sigma_\beta A^{-1\beta}{}_\alpha$$

If we let $\underline{\sigma}$ denote the row matrix corresponding to the components of $\sigma$ in the original basis, and $\underline{\sigma}'$ those in the new basis, then in matrix notation

$$\underline{\sigma}' = \underline{\sigma}\,\underline{A}^{-1},$$

$$\underline{\sigma} = \underline{\sigma}'\underline{A}, \qquad \text{or} \quad \sigma_\alpha = \sigma_\beta A^\beta{}_\alpha.$$

Thus index positioning also indicates whether the index transforms by the matrix $\underline{A}$ or its inverse $\underline{A}^{-1}$.

For example, suppose $C : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation with matrix $\underline{C} = (C^\alpha{}_\beta)$:

$$C(\underline{X}) = C^\alpha{}_\beta \underline{X}^\beta e_\alpha = C'^\alpha{}_\beta \underline{X}'^\beta e'_\alpha$$

$$C'^\alpha{}_\beta = \omega'^\alpha(C(e'_\beta)) = A^\alpha{}_\gamma\, \omega^\gamma(C(A^{-1\delta}{}_\beta e_\delta))$$

$$= A^\alpha{}_\gamma\, \omega^\gamma(C(e_\delta))\, A^{-1\delta}{}_\beta = A^\alpha{}_\gamma C^\gamma{}_\delta A^{-1\delta}{}_\beta$$

or in matrix form $\qquad \underline{C}' = \underline{A}\,\underline{C}\,\underline{A}^{-1}.$

The index notation leads to upper indices transforming by $\underline{A}$ and lower indices by $\underline{A}^{-1}$.

On the other hand suppose we evaluate the dot product in the new basis.

$$X\cdot Y = \delta_{\alpha\beta} X^\alpha Y^\beta = \delta_{\alpha\beta} A^{-1\alpha}{}_\gamma A^{-1\beta}{}_\delta X'^\gamma X'^\delta$$

$$= (X'^\alpha e'_\alpha)\cdot(Y'^\beta e'_\beta) = \underbrace{e'_\alpha \cdot e'_\beta}_{\equiv\, g'_{\alpha\beta}} X'^\alpha Y'^\beta = g'_{\alpha\beta} X'^\alpha Y'^\beta$$

i.e. $$g'_{\alpha\beta} = \delta_{\gamma\delta} A^{-1\gamma}{}_{\alpha} A^{-1\delta}{}_{\beta} = A^{-1\gamma}{}_{\alpha}\,\delta_{\gamma\delta}\,A^{-1\delta}{}_{\beta}$$

or $$\underline{g}' = (\underline{A}^{-1})^T \underline{1}\,\underline{A}^{-1}$$

where the transpose is necessary so ~~the~~ a column index is summed against a row index to make matrix multiplication.

Although the matrix of a linear transformation and the matrix of components of the dot product are both matrices, they transform differently and again the index convention makes the distinction automatically. Notice also that the relation

$$X_\alpha = \delta_{\alpha\beta} X^\beta$$

transforms to $X'_\alpha = g'_{\alpha\beta} X'^\beta$, so the correspondence between a vector and the associated covector is no longer in general simple.

The dot product is a realvalued function of a pair of vectors which is linear in each argument, hence called bi-linear. Suppose we introduce the real valued function of a ~~ve~~ covector and a vector by

$$E(\sigma, \underline{X}) = \sigma(\underline{X}). = \sigma_\alpha X^\alpha = \sigma_\alpha \delta^\alpha{}_\beta X^\beta$$

This is also ~~bilin~~ linear in each argument, and is just the evaluation function whose components are the "mixed" Kronecker delta symbol. In general we are being led to multilinear functions of a certain number of covectors and vectors. These are called tensors over $\mathbb{R}^n$ and to discuss them properly we need appropriate notation.

Before going on, though, notice that the only connection of the symbols $\{e_\alpha\}$ with $\mathbb{R}^n$ was in the original definition which didn't really enter into any of the ~~for~~ subsequent formulas. In other words, given an n-dimensional vector space V with a basis $\{e_\alpha\}$, we could have carried out the entire discussion unchanged. Instead of the ~~st~~standard dot product on $\mathbb{R}^n$, we would

Euclidean signature

be dealing with a particular inner product on $V$ such that $\{e_\alpha\}$ is an orthonormal basis. If instead we relax that condition, then

$$g_{\alpha\beta} = e_\alpha \cdot e_\beta \quad , \quad \cancel{X \cdot Y = g_{\alpha\beta} X^\alpha Y^\beta}$$

in fact defines an inner product on $V$, if $\underline{g}$ is a symmetric matrix, ~~It will be nondegerate if $\det \underline{g} \neq 0$~~. by extending it to all vectors by linearity:

$$X \cdot Y = (X^\alpha e_\alpha) \cdot (Y^\beta e_\beta) = g_{\alpha\beta} X^\alpha Y^\beta ,$$

It will be nondegenerate if $\det \underline{g} \neq 0$.

Thus our discussion goes over intact to an $n$-dimensional vector space $V$ with the addition structure of an inner product. Before we continue from such a more general viewpoint, let's re-examine the determinant and its interpretation.

Suppose $A: \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation. The determinant of its matrix $\underline{A} = (A^\alpha{}_\beta)$

$$\det \underline{A} = \epsilon_{\alpha_1 \cdots \alpha_n} A^{\alpha_1}{}_1 \cdots A^{\alpha_n}{}_n = \epsilon^{\alpha_1 \cdots \alpha_n} A^1{}_{\alpha_1} \cdots A^n{}_{\alpha_n}$$

now requires a raised alternating $\epsilon$-symbol for consistency with the index convention, numerically equal to the lower $\epsilon$-symbol. To interpret the determinant, let $X_{(\beta)} = A(e_\beta) = A^\alpha{}_\beta e_\alpha$ be the images of the basis vectors, i.e., $X^\alpha{}_{(\beta)} \equiv A^\alpha{}_\beta$, or equivalently the set of vectors corresponding to the column vectors of the matrix $\underline{A}$.

Define

$$\det(X_{(1)}, \ldots, X_{(n)}) = \det(X^\alpha{}_{(\beta)}) = \epsilon_{\alpha_1 \cdots \alpha_n} X^{\alpha_1}{}_{(1)} \cdots X^{\alpha_n}{}_{(n)} ,$$

re-interpreting the determinant as a function with $n$ vector arguments which is clearly linear in each argument alone. [Fixing the rest of the arguments, the determinant is just a certain linear combination of the components of a given argument, i.e., a covector.]

$$\det(\ldots, aX + bY, \ldots) = a \det(\ldots, X, \ldots) + b \det(\ldots, Y, \ldots).$$

We now show that the familiar properties of the determinant make its absolute value measures the volume of the parallelopiped formed by the $n$ vectors which are its argument.

(1) multiplying a column/row by a real number multiplies the determinant by the same number [ Follows from scalar multiple specialization of linearity ]

(2) interchanging two columns/rows changes the sign of the determinant (hence the determinant of a matrix with two identical or by (1) merely proportional columns/rows is zero).

$\cancel{\epsilon_{\alpha_1 \ldots \alpha_i \alpha_j}}$

$$\det(X_{(1)}, \ldots X_{(j)}, \ldots, X_{(i)}, \ldots) = \epsilon_{\alpha_1 \ldots \alpha_i \ldots \alpha_j \ldots} X^{\alpha_1}_{(1)} \cdots X^{\alpha_i}_{(j)} \cdots X^{\alpha_j}_{(i)} \cdots \quad \text{(def)}$$

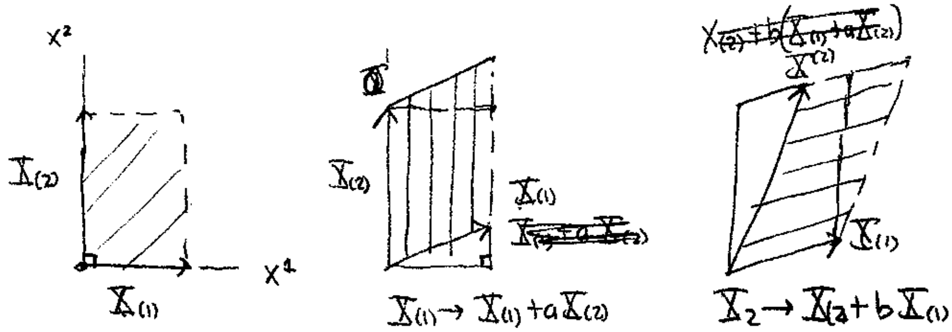
$$= -\epsilon_{\alpha_1 \ldots \alpha_j \ldots \alpha_i \ldots} X^{\alpha_1}_{(1)} \cdots X^{\alpha_i}_{(j)} \cdots X^{\alpha_j}_{(i)} \cdots \quad \text{transposition changes sign of permutation}$$

$$= -\epsilon_{\alpha_1 \ldots \alpha_i \ldots \alpha_j \ldots} X^{\alpha_1}_{(1)} \cdots X^{\alpha_j}_{(j)} \cdots X^{\alpha_i}_{(i)} \cdots \quad \text{dummy index exchange}$$

$$= -\epsilon_{\alpha_1 \ldots \alpha_i \ldots \alpha_j \ldots} X^{\alpha_1}_{(1)} \cdots X^{\alpha_i}_{(i)} \cdots X^{\alpha_j}_{(j)} \cdots \quad \text{commutivity of real numbers}$$

$$= -\det(X_{(1)}, \ldots, X_{(i)}, \ldots, X_{(j)}, \ldots) \quad \text{definition.}$$

(3) adding a real multiple of another column/row to a column/row doesnt change the determinant (by linearity plus (2))




Start with $n=2$. Take $X_{(1)}$ along $e_1$ and $X_{(2)}$ along $e_2$:

$$\det(X_{(1)}, X_{(2)}) = \epsilon_{12} X^1_{(1)} X^2_{(2)} \quad \text{(only nonzero term)}$$
$$= X^1_{(1)} X^2_{(2)} \quad \text{so} \quad |\det(X_{(1)}, X_{(2)})| = \|X_{(1)}\| \, \|X_{(2)}\| =$$
$$= \text{vol}(X_{(1)}, X_{(2)})$$

Now adding a multiple of one vector to another clearly doesnt change the area since triangular regions are simply translated, so one can move $X_{(1)}$ and $X_{(2)}$ to general positions (plus permutations for some special positions) without changing either the area or the determinant. Thus starting from our notion of area as a product of orthogonal lengths, represented here by the orthonormal basis $\{e_1, e_2\}$, the determinant of the matrix of components of 2 vectors with respect to this basis, in absolute value, clearly measures the volume of the parallellogram formed by two nonorthogonal vectors.

This same argument may be repeated for parallellopipeds in $n=3$ (familiar) and for $\mathbb{R}^n$ in general. Determinant plus an inner product give $n$-volume or $n$-measure in $\mathbb{R}^n$.

What is the interpretation of $\text{sgn} \det(X_{(1)}, \dots, X_{(n)})$ ?
This specifies the "orientation" of the ordered set of $n$ vectors relative to the orientation of the standard basis.

In $\mathbb{R}^2$, $e_2$ is related by a counterclockwise rotation from $e_1$, called the positive orientation. Two vectors $\{X_{(1)}, X_{(2)}\}$ with $\det(X_{(1)}, X_{(2)}) > 0$ are related in the same way (positively oriented set) but if $\det(X_{(1)}, X_{(2)}) < 0$, then $X_{(2)}$ is related by a clockwise rotation from $X_{(1)}$ (negatively oriented set). Permuting the set changes the sign of the orientation by the sign of the permutation (a transposition for $n=2$).

In $\mathbb{R}^3$ the right hand screw rule states that curling the fingers of the right hand from $e_1$ to $e_2$ leads the thumb to point in the direction of $e_3$ for the positively oriented standard basis of $\mathbb{R}^3$. A left handed frame (negatively oriented) would have $e_3$ in the opposite direction. For a general set of three linearly independent vectors, $\{X_{(a)}\}$, if $X_{(3)}$ is on the same/opposite side of the plane of $X_{(1)}$ and $X_{(2)}$ determined by the right hand rule (curl fingers from $X_{(1)}$ to $X_{(2)}$), the set is positively/negatively oriented.

In general one declares the standard basis of $\mathbb{R}^n$ to be positively oriented, and any ordered set of linearly independent vectors has the orientation determined by the sign of its determinant. Thus the determinant of the matrix of components of the ordered set with respect to the oriented, orthonormal standard basis measures the volume and if nonzero the orientation of the set. If we carry over the structure of $\mathbb{R}^n$ used so far to an arbitrary n-dimensional vector space, $V$, we need an inner product plus an orientation to discuss lengths and oriented volumes.

Last detail. What is the interpretation of the formula

$$\det \underline{A} = \epsilon^{\alpha_1\cdots\alpha_n} A^1{}_{\alpha_1}\cdots A^n{}_{\alpha_n} \qquad ?$$

Given any linear map $\quad A : \mathbb{R}^n \to \mathbb{R}^m \quad$ with

$$Y = A(X) \;\to\; Y^i = A^i{}_\beta X^\beta \;\to\; \underline{Y} = \underline{A}\,\underline{X},$$

one can always introduce the transpose map

$A^T : \mathbb{R}^{m*} \to \mathbb{R}^{n*}$

from the image space dual space back to the domain space dual space by defining

$$A^T(\Theta)(X) = \Theta(A(X)) \qquad \text{for } X \in \mathbb{R}^n,\ \Theta \in \mathbb{R}^{m*}$$

$$= \underbrace{\Theta_i A^i{}_\beta}_{[A^T(\Theta)]_\beta} X^\beta$$

$$\equiv \sigma_\beta$$

or $\quad \underline{\sigma} = \underline{\theta}\,\underline{A} \quad$ if we use row matrices to represent the components of the covectors. If we return to the all lower index symbol convention we would write this $\qquad \sigma_i = A_{ji}\,\theta_i$

so treating the covectors as column matrices, the matrix of $A^T$ comes out to be $\underline{A}^T$, the transpose of $\underline{A}$.

For $m=n$, we can define $\cancel{\sigma^{(\beta)} = A^\beta{}_\alpha}$

$n$ covectors $\sigma^{(\beta)}$ by $\quad \sigma^{(\beta)}{}_\alpha = A^\beta{}_\alpha \quad$ corresponding to the row vectors of the matrix $\underline{A}$ (column vectors of $\underline{A}^T$), which are just the images of the dual basis vectors under the transpose map $\qquad \sigma^{(\beta)} = A^\beta{}_\alpha\,\omega^\alpha = A^T(\omega^\beta)$

since $\qquad A^T(\omega^\beta)(e_\alpha) = \omega^\beta(A(e_\alpha)) = A^\gamma{}_\alpha\,\omega^\beta(e_\gamma) = A^\beta{}_\alpha$.

Then $\quad \det \underline{A} = \epsilon^{\alpha_1\cdots\alpha_n}\,\sigma^{(1)}{}_{\alpha_1}\cdots\sigma^{(n)}{}_{\alpha_n}$

may be interpreted as a multilinear function of $n$ covectors.

This too can be related to volumes once we have an inner product for covectors. Suppose we declare $\{\omega^\alpha\}$ to be an orthonormal basis for a dot product of covectors:

$$\omega^\alpha \cdot \omega^\beta = \delta^{\alpha\beta} \quad , \qquad \sigma \cdot \theta = \delta^{\alpha\beta}\sigma_\alpha\theta_\beta \, .$$

Then a parallel discussion can be made regarding volumes & orientations in the dual space. Again the index convention distinguishes different algebraic concepts which the usual elementary algebra notation obscures.

In $\mathbb{R}^n$ not only do we have the notion of lengths and volumes associated with 1 and n dimensional subspaces, but also p-measures associated with all possible p-dimensional subspaces $1 \le p \le n$. In $\mathbb{R}^3$ there are areas of parallelograms, namely 2-parallellopipeds, in addition to ordinary parallellopipeds, or 3-parallellopipeds. We need to generalize the determinant to discuss this question.