ULRICH THIEL
ulthiel.com/math
May 13, 2024

# COVARIANCE AND CONTRAVARIANCE (OF VECTORS)

Let $V$ be an $n$-dimensional vector space over a field $k$.
A choice of basis $e_1, \ldots, e_n$ of $V$ yields a vector space isomorphism $\psi : V \longrightarrow K^n$ sending $e_i$ to the $i$-th unit vector $(0,..,0,1,0,\ldots,0)^T$ in $k^n$. When expressing an element $v \in V$ in the basis as

$$v = \sum_{j=1}^{n} v^j e_j \quad , \quad v^j \in k ,$$

then

$$\psi(v) = (v^1, .., v^n)^T \in k^n$$

is the coefficient vector of $v$ in the basis $e_1, \ldots, e_n$.

Now, the choice of basis on $V$ may induce a natural choice of basis on other vector spaces built from $V$, like the dual $V^*$: the linear forms $e^1, .., e^n \in V^*$ uniquely defined by

$$e^i(e_j) = \delta_{ij}$$

form a basis of $V^*$. We thus also have coefficient vectors for elements of $V^*$.

An element of $K^n$ does not "know" whether it comes from $V$ or from $V^*$. But there is an important difference nonetheless! Namely, a change of basis on $V$ induces a change of basis on $V^*$, and the coefficient vectors change according to different rules!

Explicitly, let $e_1', \ldots, e_n'$ be another basis on $V$. Then there is a change of basis transformation $\phi : V \to V$, $e_i \mapsto e_i'$.

Write
$$e_j' = \phi(e_j) = \sum_{i=1}^{n} A_{ij} e_i \quad , \quad A_{ij} \in k.$$
Then $A = (A_{ij})$ is the matrix of $\phi$ in the basis $e_1, .., e_n$.

Write
$$\phi^{-1}(e_j) = \sum_{i=1}^{n} B_{ij} e_i \quad , \quad B_{ij} \in k.$$
Then $B = (B_{ij})$ is the matrix of $\phi^{-1}$ in the basis $e_1, .., e_n$.

We have
$$\begin{aligned} e_j &= \phi(\phi^{-1}(e_j)) = \phi\Big(\sum_{k=1}^{n} B_{kj} e_k\Big) = \sum_{k=1}^{n} B_{kj}\, \phi(e_k) = \sum_{k=1}^{n} B_{kj}\, e_k' \\ &= \sum_{k=1}^{n} B_{kj} \sum_{i=1}^{n} A_{ik} e_i = \sum_{i=1}^{n} \Big( \sum_{k=1}^{n} A_{ik} B_{kj} \Big) e_i \\ &= \sum_{i=1}^{n} (AB)_{ij}\, e_i \implies AB = 1 \end{aligned}$$
It follows that $A$ is invertible and $B = A^{-1}$, so $A^{-1}$ is the matrix of $\phi^{-1}$ in the basis $e_1, .., e_n$.

Let $v \in V$. Write
$$v = \sum_{j=1}^{n} v^j e_j = \sum_{j=1}^{n} v^{j'} e_j' \quad , \quad v^j, v^{j'} \in k.$$
We have two coefficient vectors $(v^1, .., v^n)^T$ and $(v^{1'}, .., v^{n'})^T$ of $v$.

How do they relate? We have
$$v = \sum_{j=1}^{n} v^j e_j = \sum_{j=1}^{n} v^j \Big( \sum_{k=1}^{n} B_{kj} e_k' \Big) = \sum_{k=1}^{n} \Big( \sum_{j=1}^{n} B_{kj} v^j \Big) e_k' ,$$
hence
$$(v^{1'}, .., v^{n'})^T = B \cdot (v^1, .., v^n)^T.$$

Hence, when changing the basis of $V$ with transformation matrix $A$, then the coefficient vector of elements of $V$ are transformed using the inverse matrix $A^{-1}$.

This behavior of the coefficient vectors is called contravariance: the coefficients contra-vary with a change of basis.

Now, consider the dual space $V^*$. Let $e^{1'},..,e^{n'}$ be the dual basis of $e_1',...,e_n'$. The map $\phi: V \to V^*$ induces a dual map $\phi^*: V^* \to V^*$ defined by

$$f \longmapsto f \circ \phi .$$

Note:

$$\phi^*(e^{i'})(e_j) = e^{i'}(\phi(e_j)) = e^{i'}(e_j') = \delta_{ij},$$

so

$$\phi^*(e^{i'}) = e^i .$$

Hence the transformation on $V^*$ sending the basis $e^1,..,e^n$ to the basis $e^{1'},..,e^{n'}$ is $(\phi^*)^{-1}$. What is the matrix of this transformation?

Note that if $f \in V^*$ with $f = \sum_{j=1}^n f_j e^j$ then

$$f(e_j) = \sum_{i=1}^n f_i e^i(e_j) = \sum_{i=1}^n f_i \delta_{ij} = f_j ,$$

We have

$$\begin{aligned}\phi^*(e^i)(e_j) &= e^i(\phi(e_j)) = e^i(e_j') = e^i\left(\sum_{k=1}^n A_{kj} e_k\right) \\ &= \sum_{k=1}^n A_{kj} e^i(e_k) = \sum_{k=1}^n A_{kj} \delta_{ik} \\ &= A_{ij}\end{aligned}$$

Hence,

$$\phi^*(e^i) = \sum_{j=1}^{n} A_{ij} e^j .$$

It follows that $A^T$ is the matrix of $\phi^*$ in the basis $e^1, \dots, e^n$.

As argued above for $V$, it follows that $(A^T)^{-1}$ is the matrix of $(\phi^*)^{-1}$ in the basis $e^1, \dots, e^n$. Consequently, writing $f \in V^*$ as

$$f = \sum_{j=1}^{n} f_j e^j = \sum_{j=1}^{n} f_j' e^{j'},$$

the two coefficient vectors $(f_1, \dots, f_n)^T$ and $(f_1', \dots, f_n')^T$ are related by

$$(f_1', \dots, f_n')^T = A^T \cdot (f_1, \dots, f_n)^T.$$

This behavior of the coefficient vectors is called covariance.

One often writes the coefficient vectors of elements of $V^*$ as row vectors and then the above becomes

$$(f_1', \dots, f_n') = (f_1, \dots, f_n) \cdot A ,$$

getting rid of the transpose for the transformation matrix.

To distinguish, coefficient vectors of elements of $V$ are called vectors and coefficient vectors of elements of $V^*$ are called covectors.

By convention one uses upper indices for vectors and lower indices for covectors.

Also, one uses the Einstein summation convention: if an index appears twice in a term, one needs to sum over all possible indices. This is just for simplifying expressions.

So,

$$\sum_{j=1}^{n} v^j e_j \text{ becomes } v^j e_j \;;\; \sum_{j=1}^{n} f_j e^j \text{ becomes } f^j e_j.$$

More generally, a choice of basis on $V$ induces a natural choice of basis on $V_{m,n} := (V^*)^{\otimes m} \otimes V^{\otimes n}$ in a change of basis on $V$ induces a change of basis on $V_{m,n}$. To emphasize the transformation rule, one says that coefficient vectors $(\in K^{m \cdot n})$ of elements of $V_{m,n}$ are of type $(m,n)$.