Tensor Tricks

Created January 2020, George Keeling

This gives reasons for the tensor tricks in Commentary Important Equations. We start with an error by Carroll.

Contents

What tensor rank?

Early in the book (page 21) we are told that "a tensor T of type (or rank) (k,l) is a multilinear map from a collection of dual vectors and vectors to $\mathbf{R}$:

$$T: \left(T_p^* \times \cdots \times T_p^*\right)_{k\text{ times}} \times \left(T_p \times \cdots \times T_p\right)_{l\text{ times}} \rightarrow \mathbf{R} \tag{1}$$

That is Carroll's 1.56 and I am pretty sure he has that the wrong way round. Dual vectors (aka covariant vectors, covectors, one forms) are written with the * and the index is downstairs. In this case the p is not that kind of index. Vectors (aka contravariant vectors) are written without the * and the index is upstairs and k is the number of up indices. So it should say "a tensor T of type (or rank) (k, l) is a multilinear map from a collection of k vectors and l dual vectors to $\mathbf{R}$." They should be rewritten as in #3 of the PF post (where dual vectors are called linear forms!).

I could not see the point of saying how many parts of the different kinds a tensor was composed of and I suggested on PF that one only needs the total $k + l$ because one can always lower or raise indices with the metric tensor $g_{\mu\nu}$ (which Carrol tells us is a $(0,2)$ tensor). This raised howls of protest from the mathematicians. It was cleared up by Ibix in #5 who wrote " in general relativity (which I think is your main interest here) you always have a metric available on any manifold of interest. That means that you can always raise or lower an index, so there isn't really any extra information in a vector that isn't encoded in its dual."

In GR, we only need to know the total rank, $k + l$, of a tensor.

The implication of Ibix's post is that there are manifolds without metrics. I never go there.

Physics Forums post.

Multi-dimensional Chain Rule

If we have three manifolds, not necessarily of the same dimension, with coordinates x^a, y^b, z^c and

$$y^b = f^b(x^a) \ , \qquad z^c = g^c(y^b) \tag{2}$$

and

$$z^c = (g \circ f)^c(x^a) = g^c\left(f^b(x^a)\right) \tag{3}$$

then

$$\frac{\partial}{\partial x^a}(g \circ f)^c = \sum_b \frac{\partial f^b}{\partial x^a}\frac{\partial g^c}{\partial y^b} \tag{4}$$

This is the same as the much more useful

$$\boxed{\frac{\partial x^c}{\partial x^a} = \frac{\partial x^b}{\partial x^a}\frac{\partial x^c}{\partial x^b}} \tag{5}$$

(x^i are usually different coordinate systems.)
Also see Commentary 2.2 Chain rule in multiple dimensions.pdf
Also in that document, why

$$\frac{\partial x^a}{\partial x^b}\frac{\partial x^c}{\partial x^d} \neq \frac{\partial x^a}{\partial x^d}\frac{\partial x^c}{\partial x^b} \tag{6}$$

Partial derivative of components gives Kronecker delta

Coordinates

x^i must be the same coordinate systems.

$$\frac{\partial x^a}{\partial x^b} = \delta^a_b \tag{7}$$

It gets better, if $a \neq b$:

$$\frac{\partial}{\partial x^b} f(x^a) = 0 \tag{8}$$

this is surprisingly easy to forget.

Vectors (C1.152)

At his 1.152 Carroll uses "the general rule, for any object with one index such as V_μ, that

$$\frac{\partial V_a}{\partial V_b} = \delta_b^a \tag{9}$$

because each component of V_μ is treated as a distinct variable."

Tensors

As I found in Commentary 4.1 Voss-Weyl formula and more.pdf this can be extended to tensors:

$$\frac{\partial T_{ik}}{\partial T_{ij}} = \delta_j^k \tag{10}$$

Partial derivatives commute

x^i must be the same coordinate systems.

$$\frac{\partial^2}{\partial x^a \partial x^b} = \frac{\partial^2}{\partial x^b \partial x^a} \tag{11}$$

See Question 002 Second order partial derivatives vanish.docx for how to go wrong!

Metric is always symmetric (C section 2.5)

$$g_{\mu\nu} = g_{\nu\mu} \tag{12}$$

Contracting with metric lowers / raises index

$$g_{\mu\nu} U^\nu = U_\mu \ , \ \ g^{\mu\nu} U_\mu = U^\nu \tag{13}$$

You can lower or raise indices on a tensor equation

For example lower λ in

$$S^{\lambda\kappa}{}_{\mu\nu} = wT^{\lambda\kappa} U_\mu P_\nu \Rightarrow S_\lambda{}^\kappa{}_{\mu\nu} = wT_\lambda{}^\kappa U_\mu P_\nu \tag{14}$$

and it can be applied to as many or as few indices as you like. If an index is summed over then both must go up/down: $S^{\lambda\kappa}{}_{\lambda\nu} = S_\lambda{}^{\kappa\lambda}{}_\nu$.

Do it by lowering the index you want then contracting out the metric tensor that appears:

$$S^{\lambda\kappa}{}_{\mu\nu} = wT^{\lambda\kappa} U_\mu P_\nu$$

$$g^{\lambda\tau} S_\tau{}^\kappa{}_{\mu\nu} = g^{\lambda\tau} wT_\tau{}^\kappa U_\mu P_\nu$$

$$g_{\lambda\rho} g^{\lambda\tau} S_\tau{}^\kappa{}_{\mu\nu} = g_{\lambda\rho} g^{\lambda\tau} wT_\tau{}^\kappa U_\mu P_\nu$$

$$\delta_\rho^\tau S_\tau{}^\kappa{}_{\mu\nu} = \delta_\rho^\tau wT_\tau{}^\kappa U_\mu P_\nu$$

$$S_\rho{}^\kappa{}_{\mu\nu} = wT_\rho{}^\kappa U_\mu P_\nu$$

Where indices are summed over, it is simpler: $S^{\lambda\kappa}{}_{\lambda\nu} = g^{\lambda\rho} g_{\sigma\lambda} S_\rho{}^{\kappa\sigma}{}_\nu = \delta_\sigma^\rho S_\rho{}^{\kappa\sigma}{}_\nu = S_\rho{}^{\kappa\rho}{}_\nu$

Swap indices with metric or any similar tensor

If both indices on the metric (or inverse metric) are being contracted with another expression, then those two indices on the other expression can be swapped. For example

$$g_{\mu\nu}\frac{\partial x^{\nu}}{\partial x^{\alpha}}\frac{\partial^2 x^{\mu}}{\partial x^{\beta'}\partial x^{\gamma}} = g_{\mu\nu}\frac{\partial x^{\mu}}{\partial x^{\alpha}}\frac{\partial^2 x^{\nu}}{\partial x^{\beta'}\partial x^{\gamma}} \tag{15}$$

The reason is that we can first swap the dummy variables μ, ν and then, because $g_{\mu\nu}$ is symmetric, they can be swapped back on g only.
This would also apply to any symmetric tensor.

Inverses and determinants

(16)-(22) are proved or discussed more fully in Commentary 2.8 Inverses Levi-Civita and more.pdf. $|A|$ is the determinant of A and may be negative. $\bar{A}^{ij}$ is the inverse of A_{ij} so $\bar{A}^{ij}A_{jk} = \delta^i_k$ and $\bar{A}^{ij} \neq A^{ij}$ unless A is the metric. The determinant of the metric is often written just g instead of $|g|$. ϵ is the Levi-Civita number.

Inverse of a matrix

For a square matrix A and its inverse $\bar{A}$

$$\bar{A} = \frac{1}{|A|}\mathrm{adj}(A) \tag{16}$$

The adjugate is the transpose of the cofactor matrix. The cofactor matrix is the matrix whose i, jth element is the $(-1)^{i+j}$ times the determinant of the matrix formed by removing row i column j from A.

The determinant of the inverse is reciprocal of the determinant

$$|\bar{A}| = \frac{1}{|A|} \tag{17}$$

Determinant of a tensor in terms of Levi-Civita symbol (C Eq 2.66)

Carroll's 2.66 for the determinant of a matrix A with components $A^{\mu_1}{}_{\mu'_1}$, is

$$\epsilon_{\mu'_1\mu'_2\dots\mu'_n}|A| = \epsilon_{\mu_1\mu_2\dots\mu_n}A^{\mu_1}{}_{\mu'_1}A^{\mu_2}{}_{\mu'_2}\dots A^{\mu_n}{}_{\mu'_n} \tag{18}$$

That is $n!$ ways of calculating it and we can select one by setting $\mu'_1\mu'_2\dots\mu'_n = 12\dots n$ (assuming we are indexing from 1). For a fully covariant tensor that would be

$$|A| = \epsilon^{\mu_1\mu_2\dots\mu_n}A_{\mu_1 1}A_{\mu_2 2}\cdots A_{\mu_n n} \tag{19}$$

Inverse tensor

If $\bar{A}^{\mu\nu}$ is the inverse of a tensor $A_{\mu\nu}$ and $|A|$ is its determinant then, the inverse is given by

$$|A|\bar{A}^{\alpha\mu_\alpha} = \epsilon^{\mu_1\mu_2\mu_3\dots\mu_n}\prod_{\beta\neq\alpha}A_{\mu_\beta\beta} \tag{20}$$

I got three likes for this formula on Physics forums! The discussion started about here.

A relationship for the derivative of the determinant

$$|\bar{A}|\partial_\nu |A| = \bar{A}^{\rho\sigma}\partial_\nu A_{\sigma\rho} \text{ or } \partial_\nu |A| = |A|\bar{A}^{\rho\sigma}\partial_\nu A_{\sigma\rho} \tag{21}$$

The former is more beautiful, the latter more useful. They are the same due to (17). For the metric the latter becomes

$$\partial_\mu g = g g^{\nu\lambda}\partial_\mu g_{\nu\lambda} \tag{22}$$

Fully contracted symmetric × antisymmetric tensor vanishes

If $A^{\mu\nu}$ is antisymmetric and $S_{\mu\nu}$ is symmetric then

$$A^{\mu\nu}S_{\mu\nu} = 0 \tag{23}$$

By definition $A^{\mu\nu} = -A^{\nu\mu}$ therefore

$$-A^{\nu\mu}S_{\mu\nu} = A^{\mu\nu}S_{\mu\nu} \tag{24}$$

$$A^{\mu\nu}S_{\mu\nu} = A^{\nu\mu}S_{\nu\mu} \text{ changing dummy indices} \tag{25}$$

$$A^{\nu\mu}S_{\nu\mu} = A^{\nu\mu}S_{\mu\nu} \text{ by symmetry of } S \tag{26}$$

Therefore

$$-A^{\nu\mu}S_{\mu\nu} = A^{\nu\mu}S_{\mu\nu} \tag{27}$$

and the only solution to that is (23).

Symmetrising a tensor equation

If the same vector occurs twice in an equation with different dummy indices, the equation can be symmetrised more. Example:

$$p^\mu p^\nu \nabla_\mu K_\nu = p^\mu p^\nu \nabla_{(\mu}K_{\nu)} \tag{28}$$

Proof

$$p^\mu p^\nu \nabla_{(\mu}K_{\nu)} = \frac{1}{2}\left(p^\mu p^\nu \nabla_\mu K_\nu + p^\mu p^\nu \nabla_\nu K_\mu\right) \tag{100}$$

$$= \frac{1}{2}\left(p^\mu p^\nu \nabla_\mu K_\nu + p^\nu p^\mu \nabla_\mu K_\nu\right) = p^\mu p^\nu \nabla_\mu K_\nu \tag{101}$$

Two formulas involving four-velocity

The four velocity is defined as $U^\mu = dx^\mu/d\tau$ then, for timelike paths, we have

$$\boxed{U_\nu U^\nu = -1} \tag{29}$$

Proof:
By definition

$$U_\nu U^\nu = g_{\mu\nu}U^\mu U^\nu = g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} \tag{30}$$

We have the metric / line element equation with proper time τ that

$$ds^2 = -d\tau^2 = g_{\mu\nu}dx^\mu dx^\nu \tag{31}$$

$$\Rightarrow -1 = g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = g_{\mu\nu}U^\mu U^\nu = U_\nu U^\nu \tag{32}$$

For null paths, $d\tau = 0$, so that does not work.

Second formula

$$U_\nu \nabla_\mu U^\nu = 0 \tag{33}$$

and by (14) $U^\nu \nabla_\mu U_\nu = 0$.

Proof:

$$\nabla_\mu(U_\alpha U^\beta) = U_\alpha \nabla_\mu U^\beta + U^\beta \nabla_\mu U_\alpha \tag{34}$$

$$= g_{\alpha\nu} U^\nu \nabla_\mu (g^{\beta\lambda} U_\lambda) + U^\beta \nabla_\mu U_\alpha \tag{35}$$

$$= g_{\alpha\nu} g^{\beta\lambda} U^\nu \nabla_\mu U_\lambda + U^\beta \nabla_\mu U_\alpha \tag{36}$$

$$\Rightarrow \nabla_\mu(U_\rho U^\rho) = g_{\rho\nu} g^{\rho\lambda} U^\nu \nabla_\mu U_\lambda + U^\rho \nabla_\mu U_\rho = \delta^\lambda_\nu U^\nu \nabla_\mu U_\lambda + U^\rho \nabla_\mu U_\rho \tag{37}$$

$$= U^\lambda \nabla_\mu U_\lambda + U^\rho \nabla_\mu U_\rho = 2U^\rho \nabla_\mu U_\rho \tag{38}$$

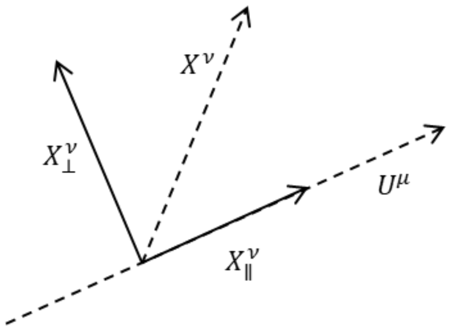
But $\nabla_\mu(U_\rho U^\rho) = 0$ therefore $U^\rho \nabla_\mu U_\rho = 0$.

The projection tensor on four-velocity (C Eq 1.121)

To project a vector X^ν orthogonal to a four velocity U^μ contract X^ν with the projection tensor

$$P^\sigma{}_\nu = \delta^\sigma_\nu + U^\sigma U_\nu \tag{39}$$

That means $P^\sigma{}_\nu X^\nu$ will be the projection of X^ν orthogonal to U^μ. We call that vector $X^\nu_\perp$.



To convince ourselves that $P^\sigma{}_\nu$ does the trick take a vector $V^\mu_\parallel$ parallel to U^μ then

$$P^\sigma{}_\nu V^\nu_\parallel = \delta^\sigma_\nu V^\nu_\parallel + U^\sigma U_\nu V^\nu_\parallel \tag{40}$$

If the two vectors are parallel then for some scalar k, we have $V^\mu_\parallel = kU^\mu$. Put that into the equation and using (29)

$$P^\sigma{}_\nu V^\nu_\parallel = V^\sigma_\parallel + kU^\sigma U_\nu U^\nu \tag{41}$$

$$\Rightarrow P^\sigma{}_\nu V^\nu_\parallel = kU^\sigma - kU^\sigma = 0 \tag{42}$$

Then take a vector $W^\mu_\perp$ perpendicular to U^μ then

$$P^\sigma{}_\nu W^\nu_\perp = \delta^\sigma_\nu W^\nu_\perp + U^\sigma U_\nu W^\nu_\perp = W^\sigma_\perp + 0 \tag{43}$$

because $U_\nu W^\nu_\perp$ is a dot product and we know it vanishes for orthogonal vectors.

Any vector X^{ν} can be written as the sum of two vectors $V_{\parallel}^{\nu}$ and $W_{\perp}^{\nu}$ so

$$P^{\sigma}{}_{\nu}X^{\nu} = P^{\sigma}{}_{\nu}\left(V_{\parallel}^{\nu} + W_{\perp}^{\nu}\right) = P^{\sigma}{}_{\nu}W_{\perp}^{\nu} \tag{44}$$

which is by definition the projection of X^{ν} orthogonally.

Contra / co-variant tensor transformation matrices

In exercise 1.10 on the Lorentz transformation we had the Lorentz transformation for contravariant components:

$$x^{\mu'} = \Lambda^{\mu'}{}_{\nu}x^{\nu} \tag{45}$$

but needed to transform the covariant form of the EM field tensor $F_{\mu\nu}$. One way to do this is to find the contravariant form $F^{\mu\nu}$ and use $\Lambda^{\mu'}{}_{\nu}$, a quicker way is to find the transformation matrix $\Lambda^{\nu}{}_{\kappa'}$ for covariant indices and use that on $F_{\mu\nu}$.

The transformation matrix for opposite type of index is simply the inverse of the transformation index:

$$\Lambda^{\mu}{}_{\mu'} = \left(\Lambda^{\mu'}{}_{\mu}\right)^{\mathrm{I}} \tag{46}$$

We have the transformation and its inverse:

$$x^{\mu'} = \Lambda^{\mu'}{}_{\nu}x^{\nu} \ , \qquad x^{\nu} = \Lambda^{\nu}{}_{\kappa'}x^{\kappa'} \tag{47}$$

Putting those together

$$x^{\mu'} = \Lambda^{\mu'}{}_{\nu}\Lambda^{\nu}{}_{\kappa'}x^{\kappa'} \tag{48}$$

Clearly the transformation and its inverse must be the identity so

$$x^{\mu'} = \Lambda^{\mu'}{}_{\nu}\Lambda^{\nu}{}_{\kappa'}x^{\kappa'} = \delta^{\mu'}_{\kappa'}x^{\kappa'} = x^{\mu'} \tag{49}$$

Which is obvious. But $\Lambda^{\nu}{}_{\kappa'}$ is the transformation for a covariant component and from (49) we have

$$\Lambda^{\mu'}{}_{\nu}\Lambda^{\nu}{}_{\kappa'} = \delta^{\mu'}_{\kappa'} \tag{50}$$

so **the transformation for covariant components is simply the inverse of the transformation for contravariant components.** That was not so obvious.

Tensor contractions using matrices

Also in exercise 1.10 we had to calculate

$$F^{\mu'\nu'} = \Lambda^{\mu'}{}_{\mu}\Lambda^{\nu'}{}_{\nu}F^{\mu\nu} \tag{51}$$

This can be done by doing the sums indicated by the μ, ν indices or by converting the expression to matrices. The latter is much simpler (e.g. $16 \times (4 \times 4)$ calculations vs 2 matrix multiplications). But care must be taken!

$$\Lambda^{\mu'}{}_{\mu}\Lambda^{\nu'}{}_{\nu}F^{\mu\nu} = \Lambda^{\mu'}{}_{\mu}F^{\mu\nu}\Lambda^{\nu'}{}_{\nu} = \Lambda^{\mu'}{}_{\mu}(F\Lambda^{T})^{\mu\nu'} \tag{52}$$

$$\Lambda^{\mu'}{}_{\mu}(F\Lambda^{T})^{\mu\nu'} = (\Lambda F\Lambda^{T})^{\mu'\nu'} \tag{53}$$

When tensors are written as matrices left to right indices of the tensor strictly correspond to rows, columns. So the order of μ', ν is important in Λ.
The contracted (summed) indices must always be next to each other to make the matrix multiplication work. Thus we must use the transpose of Λ in (52).
It is also important to keep track of the order indices of the resulting tensor as we see below.

As it happens, Λ is a transformation matrix not a tensor. It does not affect the argument.

We also had to calculate

$$F^{\mu\nu} = \eta^{\nu\sigma}\eta^{\mu\rho}F_{\rho\sigma} \tag{54}$$

$\eta^{\nu\sigma}$ is the symmetric Minkowski inverse metric. That goes

$$\eta^{\nu\sigma}\eta^{\mu\rho}F_{\rho\sigma} = \eta^{\nu\sigma}(\eta F)^{\mu}{}_{\sigma} \tag{55}$$

$$= \left(\eta\left((\eta F)^{\mu}{}_{\sigma}\right)^{T}\right)^{\nu\mu} \tag{56}$$

In (55) the contraction could be done immediately because the contracted ρ indices were next to each other. In (56) we had to take the transpose of $(\eta F)^{\mu}{}_{\sigma}$ to get the contracted σ indices next to each other. The resulting matrix has indices $\nu\mu$ in that order which is the opposite order to what they are in $F^{\mu\nu}$ so we must take the transpose of (56) again to get to $F^{\mu\nu}$.

There is a simpler way:

$$\eta^{\nu\sigma}\eta^{\mu\rho}F_{\rho\sigma} = (\eta F)^{\mu}{}_{\sigma}\eta^{\nu\sigma} = (\eta F\eta^{T})^{\mu\nu} \tag{57}$$

The indices are in the correct order and as $\eta^{\nu\sigma}$ is symmetric $\eta^{T} = \eta$.

Links to resources

My blog: https://www.general-relativity.net/
Docx: Commentary Tensor Tricks.docx

The fab four
Commentary Important Equations.pdf
Commentary 2.5 The Metric.pdf
Commentary Tensor Tricks.pdf
Commentary Constants and conversion factors.pdf