Einstein notation

From Wikipedia, the free encyclopedia

*For other topics related to **Einstein**, see Einstein (disambiguation).*

In mathematics, especially in applications of linear algebra to physics, the **Einstein notation** or **Einstein summation convention** is a notational convention useful when dealing with coordinate formulae. It was introduced by Albert Einstein in 1916 [1].

According to this convention, when an index variable appears twice in a single term, once in an upper (superscript) and once in a lower (subscript) position, it implies that we are summing over all of its possible values. In typical applications, the indices are 1,2,3 (representing the three dimensions of physical Euclidean space), or 0,1,2,3 or 1,2,3,4 (representing the four dimensions of space-time, or Minkowski space), but they can have any range, even (in some applications) an infinite set. Abstract index notation is an improvement of Einstein notation.

In general relativity, the Greek alphabet and the Roman alphabet are used to distinguish whether summing over 1,2,3 or 0,1,2,3 (usually Roman, i, j, ... for 1,2,3 and Greek, μ, ν, ... for 0,1,2,3). As in sign conventions, the convention used in practice varies: Roman and Greek may be reversed.

When there is a fixed basis, one can work with only subscripts, but in general one must distinguish between superscripts and subscripts; see below.

It is important to keep in mind that no new physical laws or ideas result from using Einstein notation; rather, it merely helps in identifying relationships and symmetries often 'hidden' by more conventional notation.

In some fields, Einstein notation is referred to simply as index notation, or indicial notation. Additionally, the use of the implied summation of repeated indices is referred to as the *Einstein Sum Convention.*

Contents

Introduction

The basic idea of Einstein notation is very simple. It allows one to replace something bulky, such as:

$$y = c_1x_1 + c_2x_2 + c_3x_3 + ... + c_nx_n$$

typically written as:

$$y = \sum_{i=1}^{n} c_i x_i$$

with something even simpler, in *Einstein notation*:

$$y = c_i x^i$$

In Einstein notation, indices such as i in the equation above can appear as either subscripts or superscripts. The position of the index has a specific meaning. It is important, of course, not to interpret an index appearing in the superscript position as if it were an exponent, which is the convention in standard algebra. Here, the superscripted i above the symbol x represents an integer-valued index running from 1 to n.

The virtue of Einstein notation is that an index appearing two or more times in a single term implies summation across that index, so that the summation symbol is unnecessary. Since the summation in effect "eliminates" the index over which the sum is taken, the summation index does not appear on the opposite side of the equals sign.

For the sequel, note that x^i is a covector on $\mathbf{R}^n$ (it takes in a vector and gives out the *i*th component), hence the upper index.

Vector representations

First, we can use Einstein notation in linear algebra to distinguish easily between vectors and covectors: upper indices represent the *components* of vectors, while lower indices represent the *components* of covectors. However, vectors themselves (not their components) have lower indices, and covectors have upper indices.

This point is frequently confused.

Given a vector space V and its dual space V^*, one represents vectors (elements of V) with subscripts, as in $v_i \in V$, and covectors with superscripts, as in $w^i \in V^*$. However, the *components* of vectors and covectors follow the opposite convention: if e_i are a basis for V and e^i are the dual basis for V^*, then vectors are represented as:

$$v = a^i e_i = \begin{bmatrix} a^1 \\ a^2 \\ \vdots \\ a^n \end{bmatrix}$$

and covectors are represented as

$$w = a_i e^i = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$$

This is because a component of a vector (a coefficient in some basis) is the *value* of a *covector*: the coefficient of e_i is the value of the corresponding covector in the dual basis: $a^i = e^i(v)$. Note that e^i is a covector, but a^i is a scalar. More prosaically, you pair components with vectors; since vectors have lower indices, components have upper indices.

In terms of covariance and contravariance of vectors, lower indices represent (components of!) covariant vectors (covectors), while upper indices represent (components of!) contravariant vectors (vectors): they transform covariantly (resp., contravariantly) with respect to change of coordinates.

A particularly confusing notation is to use the same letter both for a (co)vector and its components, as in:

$$v = v^i e_i = \begin{bmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{bmatrix}$$

$$w = w_i e^i = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix}$$

Here v^i does not mean "the covector v", but rather, "the components of the vector v".

Mnemonics

- "Upper indices go up to down; lower indices go left to right"
- You can stack vectors (column matrices) side-by-side:

$$\begin{bmatrix} v_1 & \cdots & v_k \end{bmatrix}.$$

Hence the lower index indicates which *column* you are in.

- You can stack covectors (row matrices) top-to-bottom:

$$\begin{bmatrix} w^1 \\ \vdots \\ w^k \end{bmatrix}$$

Hence the upper index indicates which *row* you are in.

Superscripts and subscripts vs. only subscripts

In the presence of a non-degenerate form (an isomorphism $V \to V^*$), (for instance a Riemannian metric or Minkowski metric), one can raise and lower indices.

A basis gives such a form (via the dual basis), hence when working on $\mathbf{R}^n$ with a fixed basis, one can work with just subscripts.

However, if one changes coordinates, the way that coefficients change depends on the variance of the object, and one cannot ignore the distinction; see covariance and contravariance of vectors.

Common operations in this notation

Inner product

Given a row vector v^i and a column vector u_i of the same size, we can take the inner product $v^i u_i$, which is a scalar: it's evaluating the covector on the vector.

Multiplication of a vector by a matrix

Given a matrix A^i_j and a (column) vector v^j, the coefficients of the product $\mathbf{A}\mathbf{v}$ are given by $A^i_j v^j$.

Matrix multiplication

We can represent matrix multiplication as:

$$C^i_k = A^i_j \cdot B^j_k$$

This expression is equivalent to the more conventional (and less compact) notation:

$$\mathbf{C} = \mathbf{A} \cdot \mathbf{B} = \sum_{j=1}^{N} A_{ij} B_{jk}$$

Trace

Given a matrix A^i_j, summing over a common index A^i_i yields the trace.

Outer product

The outer product of the column vector **u** by the row vector **v** yields an $M \times N$ matrix **A**:

$$\mathbf{A} = \mathbf{u} \cdot \mathbf{v}$$

In Einstein notation, we have:

$$A^i_j = u^i \cdot v_j = uv^i_j$$

Since *i* and *j* represent two *different* indices, and in this case over two different ranges *M* and *N* respectively, the indices are not eliminated by the multiplication. Both indices survive the multiplication to become the two indices of the newly-created matrix *A*.

Coefficients on tensors and related

Given a tensor field and a basis (of linearly independent vector fields), the coefficients of the tensor field in a basis can be computed by evaluating on a suitable combination of the basis and dual basis, and inherits the correct indexing. We list notable examples.

Throughout, let e_i be a basis of vector fields (a moving frame).

- (covariant) metric tensor

  $g_{ij} = g(e_i,e_j)$

- (contravariant) metric tensor

  $g^{ij} = g(e^i,e^j)$

- Torsion tensor (using the below)

  $$T^c_{ab} = \Gamma^c_{ab} - \Gamma^c_{ba} - \gamma^c_{ab},$$

which follows from the formula

$$T = \nabla_X Y - \nabla_Y X - [X,Y].$$

- Riemann curvature tensor

  $$R^\rho{}_{\sigma\mu\nu} = dx^\rho(R(\partial_\mu, \partial_\nu)\partial_\sigma)$$

This also applies for some operations that are not tensorial, for instance:

- Christoffel symbols

  $$\nabla_i e_j = \Gamma^k_{ij} e_k$$

where $\nabla_i e_j$ is the covariant derivative. Equivalently,

$$\Gamma^k_{ij} = e^k \nabla_i e_j$$

- commutator coefficients

$$[e_i, e_j] = \gamma_{ij}^k e_k$$

where $[e_i,e_j]$ is the Lie bracket. Equivalently,

$$\gamma_{ij}^k = e^k[e_i, e_j].$$

Vector dot product

In mechanics and engineering, vectors in 3D space are often described in relation to orthogonal unit vectors **i**, **j** and **k**.

$$\mathbf{u} = u_x\mathbf{i} + u_y\mathbf{j} + u_z\mathbf{k}$$

If the basis vectors **i**, **j**, and **k** are instead expressed as $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$, a vector can be expressed in terms of a summation:

$$\mathbf{u} = u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + u_3\mathbf{e}_3 = \sum_{i=1}^{3} u_i\mathbf{e}_i$$

In Einstein notation, the summation symbol is omitted since the index *i* is repeated and we simply write

$$\mathbf{u} = u_i\mathbf{e}_i$$

Using $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ instead of **i**, **j**, and **k**, together with Einstein notation, we obtain a concise algebraic presentation of vector and tensor equations. For example,

$$\mathbf{u}\cdot\mathbf{v} = \sum_{i=1}^{3} u_i\mathbf{e}_i \cdot \sum_{j=1}^{3} v_j\mathbf{e}_j = u_i\mathbf{e}_i \cdot v_j\mathbf{e}_j$$

or equivalently:

$$\mathbf{u}\cdot\mathbf{v} = \sum_{i=1}^{3}\sum_{j=1}^{3} u_i v_j(\mathbf{e}_i\cdot\mathbf{e}_j) = u_i v_j(\mathbf{e}_i\cdot\mathbf{e}_j)$$

where

$$\mathbf{e}_i\cdot\mathbf{e}_j = \delta_{ij}$$

and δ_{ij} is the Kronecker delta, which is equal to 1 when $i = j$, and 0 otherwise. It logically follows that this allows one *j* in the equation to be converted to an *i*, or one *i* to be converted to a *j*. Then,

$$\mathbf{u}\cdot\mathbf{v} = u^i v^j \delta_{ij} = u^i v_i = u_j v^j$$

Vector cross product

For the cross product,

$$\mathbf{u}\times\mathbf{v} = \sum_{j=1}^{3} u_j\mathbf{e}_j \times \sum_{k=1}^{3} v_k\mathbf{e}_k = u_j\mathbf{e}_j \times v_k\mathbf{e}_k = u_j v_k(\mathbf{e}_j\times\mathbf{e}_k) = \epsilon_{ijk}\mathbf{e}_i u_j v_k$$

where $\mathbf{e}_j\times\mathbf{e}_k = \epsilon_{ijk}\mathbf{e}_i$ and ϵ_{ijk} is the Levi-Civita symbol defined by:

$$\epsilon_{ijk} = \begin{cases} 0 & \text{unless } i,j,k \text{ are distinct} \\ +1 & \text{if } (i,j,k) \text{ is an even permutation of } (1,2,3) \\ -1 & \text{if } (i,j,k) \text{ is an odd permutation of } (1,2,3) \end{cases}$$

which recovers

$$\mathbf{u} \times \mathbf{v} = (u_2 v_3 - u_3 v_2)\mathbf{e}_1 + (u_3 v_1 - u_1 v_3)\mathbf{e}_2 + (u_1 v_2 - u_2 v_1)\mathbf{e}_3$$

from

$$\mathbf{u} \times \mathbf{v} = \epsilon_{ijk}\mathbf{e}_i u_j v_k = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\epsilon_{ijk}\mathbf{e}_i u_j v_k.$$

Additionally, if $\mathbf{w} = \mathbf{u} \times \mathbf{v}$, then $\mathbf{w} = \epsilon_{ijk}\mathbf{e}_i u_j v_k$ and $w_i = \epsilon_{ijk} u_j v_k$. This also highlights that when an index appears once on *both* sides of the equation, this implies a system of equations instead of a summation:

$$w_1 = \epsilon_{1jk} u_j v_k$$
$$w_2 = \epsilon_{2jk} u_j v_k$$
$$w_3 = \epsilon_{3jk} u_j v_k$$

Alternatively, this could be expressed as

$$\mathbf{u} \times \mathbf{v} = \mathbf{u} \cdot \epsilon \cdot \mathbf{v}$$

but, this isn't the notation Einstein used.

Abstract definitions

In the traditional usage, one has in mind a vector space *V* with finite dimension *n*, and a specific basis of *V*. We can write the basis vectors as $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_n$. Then if **v** is a vector in *V*, it has coordinates $v^1, \ldots, v^n$ relative to this basis.

The basic rule is:

$$\mathbf{v} = v^i \mathbf{e}_i.$$

In this expression, it was assumed that the term on the right side was to be summed as *i* goes from 1 to *n*, because the index *i* does not appear on both sides of the expression. (Or, using Einstein's convention, because the index *i* appeared twice.)

The *i* is known as a *dummy index* since the result is not dependent on it; thus we could also write, for example:

$$\mathbf{v} = v^j \mathbf{e}_j.$$

An index that is not summed over is a *free index* and should be found in each term of the equation or formula. Compare dummy indices and free indices with free variables and bound variables.

The value of the Einstein convention is that it applies to other vector spaces built from *V* using the tensor product and duality. For example, $V \otimes V$, the tensor product of *V* with itself, has a basis consisting of tensors of the form $\mathbf{e}_{ij} = \mathbf{e}_i \otimes \mathbf{e}_j$. Any tensor **T** in $V \otimes V$ can be written as:

$$\mathbf{T} = T^{ij}\mathbf{e}_{ij}.$$

*V**, the dual of *V*, has a basis $\mathbf{e}^1$, $\mathbf{e}^2$, ..., $\mathbf{e}^n$ which obeys the rule

$$\mathbf{e}^i(\mathbf{e}_j) = \delta^i_j.$$

Here δ is the Kronecker delta, so δ^i_j is 1 if $i = j$ and 0 otherwise.

As

$$\mathrm{Hom}(V, W) = V^* \otimes W$$

the row-column coordinates on a matrix correspond to the upper-lower indices on the tensor product.

Examples

Einstein summation is clarified with the help of a few simple examples. Consider four-dimensional spacetime, where indices run from 0 to 3:

$$a^\mu b_\mu = a^0 b_0 + a^1 b_1 + a^2 b_2 + a^3 b_3$$

$$a^{\mu\nu} b_\mu = a^{0\nu} b_0 + a^{1\nu} b_1 + a^{2\nu} b_2 + a^{3\nu} b_3.$$

The above example is one of contraction, a common tensor operation. The tensor $a^{\mu\nu} b_\alpha$ becomes a new tensor by summing over the first upper index and the lower index. Typically the resulting tensor is renamed with the contracted indices removed:

$$s^\nu = a^{\mu\nu} b_\mu.$$

For a familiar example, consider the dot product of two vectors **a** and **b**. The dot product is defined simply as summation over the indices of **a** and **b**:

$$\mathbf{a} \cdot \mathbf{b} = a^\alpha b_\alpha = a^0 b_0 + a^1 b_1 + a^2 b_2 + a^3 b_3,$$

which is our familiar formula for the vector dot product. Remember it is sometimes necessary to change the components of **a** in order to lower its index; however, this is not necessary in Euclidean space, or any space with a metric equal to its inverse metric (e.g., flat spacetime).

See also

- Abstract index notation
- Bra-ket notation
- Penrose graphical notation

References

- Rawlings, Steve. "Lecture 10 - Einstein Summation Convention and Vector Identities", Oxford University, 2007-02-01.

1. ^ Einstein, Albert (1916). "The Foundation of the General Theory of Relativity" (PDF). *Annalen der Physik*. Retrieved on 2006-09-03.