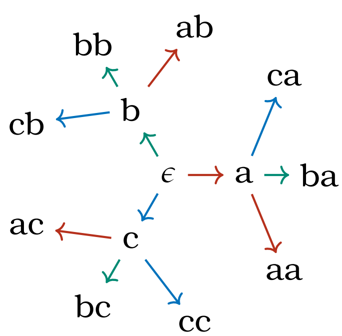
Talking Tensors

A short coarse on tensors

2024-03-23

James B. Wilson



Department of
Mathematics
Colorado State
University

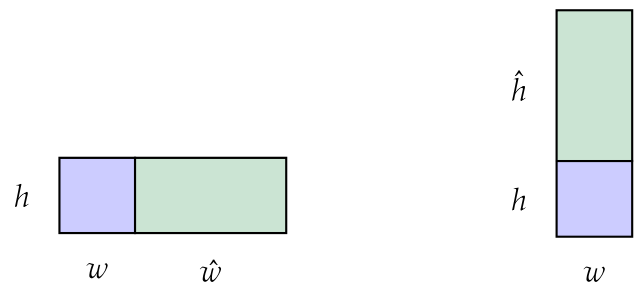
James.Wilson@ColoState.Edu

James.Wilson@ColoState.Edu

Contents

The distributive property 1

The distributive law will be the star of the tensor show. Distributive laws appear throughout our world of measurement and this offers some gentle points of entry to the study. Start with area.



So area as function $A(w,h)$ of width and height distributes over those parameters.

$$A(w+\hat{w},h) = A(w,h) + A(\hat{w},h) \qquad A(w,h+\hat{h}) = A(w,h) + A(w,\hat{h})$$

Used in repetition, all areas get defined by distributing measurement to sums of subrectangles that steadily approximate the whole. Thus area's distributive property water-marks calculus in numerous ways including the following fundamental identity.

$$\int_M (f+g)\mathbf{d}x = \int_M f\,\mathbf{d}x + \int_M g\,\mathbf{d}x$$

"Real" life

Look elsewhere in your life. When you buy x items at cost $\text{Cost}(x)$ and then buy y more the total cost is:

$$\text{Cost}(x+y) = \text{Cost}(x) + \text{Cost}(y).$$

Add the cost of tax $\text{Tax}(x)$ and the total changes to

$$(\text{Cost}+\text{Tax})(x) = \text{Cost}(x) + \text{Tax}(x).$$

Cost distributes.

Here is another life situation. The work of e employee for h hours produces $\mathrm{W}(e,h)$ widgets so at peak productivity adding $\hat{e}$ workers or $\hat{h}$ extra hours we get

$$\mathrm{W}(e+\hat{e},h) = \mathrm{W}(e,h)+\mathrm{W}(\hat{e},h) \qquad \mathrm{W}(e,h+\hat{h}) = \mathrm{W}(e,h)+\mathrm{W}(e,\hat{h}).$$

Work distributes, both over workers and over time.

Well all this is at least true for some range of values. All scientific claims of connection to reality are true only for sensible regions of values.

Need to print the names of both faculty and graduate students. You could distribute that work as well, as seen in Listing 1.0.1, the command `Print` distributes over the operator of concatenation `cat` with the result of simply appending the two individual printing commands. We could say

```
Print(as cat bs)== Print(as)append Print(bs)
```

Listing 1.0.1 Many operations in computation are distributive.

In [1]
```
fac = [ "Hamilton", "Levi-Chivita" ]
students = [ TBD use math geoneolgoy to find some
    students of these]
Print(faculty cat students)
```

Out [1]
```
Hamilton
Levi-Chivita
....TBD
```

In [2]
```
Print(faculty)
Print(students)
```

Out [2]
```
Hamilton
Levi-Chivita
....TBD
```

Now as a warning the purpose of these notes is to get deep into the theory and computation of tensors and most examples will rely on substantive mathematical background. These examples however might remind you of the connections to concrete problems and help you communicate across the sciences and perhaps even to your friends and family who wonder what you do.

Notations

By custom, area $A(\ell, w)$ of rectangle is denoted $\ell \cdot w$, called multiplication, and the distributive law takes its form:

$$(\ell + \hat{\ell}) \cdot h = \ell \cdot w + \hat{\ell} \cdot h \qquad\qquad \ell \cdot (h + \hat{h}) = \ell \cdot h + \ell \cdot \hat{h}.$$

This is the first of dozens of notations invented over the centuries to communicate functions that distribute their work amongst sums of sub-applications of the function. Most honor one of these three modes: functions on the outside applied to sums of inputs

$$f(u+\acute{u}) = f(u)+f(\acute{u}), \qquad \mathtt{Print}(students+faculty), \qquad \mathrm{Cost}(x,y,z)\ldots;$$

functions in the middle like these operators

$$(u+\acute{u}) * v = u * v + \acute{u} * v, \qquad (u+\acute{u}) \cdot v = u \cdot v + \acute{u} \cdot v, \qquad u \otimes v, \qquad uv, \ldots;$$

or functions on the outside like these

$$\langle u+\acute{u} | v\rangle = \langle u|v\rangle + \langle \acute{u}|v\rangle, \qquad [u+\acute{u}, v] = [u,v] + [\acute{u}, v], \qquad /u,v/, \ldots .$$

Axes and valence

Just, as area distributes, so does volume $V(\ell, w, h)$ only with one more axis.

[TBD: a tikz picture of three volumes showing three axes of distributive volume]

As formulas this could be written:

$$\begin{aligned}
V(\ell+\hat{\ell}, w, h) &= V(\ell, w, h) + V(\hat{\ell}, w, h)\\
V(\ell, w+\hat{w}, h) &= V(\ell, w, h) + V(\ell, \hat{w}, h)\\
V(\ell, w, h+\hat{h}) &= V(\ell, w, h) + V(\ell, w, \hat{h})
\end{aligned}$$

Venturing into hypervolumes benefits from some notation. For instance, the area of a rectangle of length ℓ and width w could be expressed as $A(w,\ell)$ or $A(\ell,w)$ without concern because it is the letter, not its position in the list, that conveys the meaning. A volume of box of length ℓ, width w and height h would similarly be $V(\ell,w,h)$ or $V(h,\ell,w)$ and so on. We can carry this convention forward by treating inputs not as ordered lists but as unordered but *labeled* lists. So we have a set A of axes an for each axis $a \in A$ the parameter in that axis is decorated by a, e.g. x_a. For example $A = \{x,y\}$ Hypervolume could then be regarded as $V(x_a \mid a \in A)$.

instead of $A(x,y)$ because of the fixed conventions of x and y axes. The letters themselves, not their position in the formula, are enough to understand the meaning. With volume $V(x,y,z)$ a similar tolerance could be had for writing $V(z,y,x)$.

To go further in It This is one of the foundational reason for the method of calculus: we can chop area up and add the resulting smaller areas with no affect on the result. It's power is reflected in the final form of calculus through properties like this:

$$\int_M (f+g)\mathrm{d}x = \int_M f\,\mathrm{d}x + \int_M g\,\mathrm{d}x$$

As area distributes, volume

1.1 Distributive algebra

What does this require?

$$u*(v+\acute{v}) = u*v + u*\acute{v} \qquad (u+\acute{u})*v = u*v+\acute{u}*v. \tag{1.1}$$

We definitely need additions, but we should not jump to assume that u, v, and w are of the same type. Just look at matrix multiplication (we use $\mathbb{R}^{a\times b}$ to denote $(a\times b)$-matrices of real numbers)

$$*:\mathbb{R}^{a\times b}\times\mathbb{R}^{b\times c}\to\mathbb{R}^{a\times c} \qquad (u*v)_{ij} = \sum_k u_{ik}v_{kj}.$$

So we except this is a study of *heterogeneous* algebra, so we wont be captivated by homomorphism but rather what we will call *heteromorphisms*. So we could think of three types of data U, V and

W each with a $+$ each combined by a function $*: U \times V \to W$ that satisfies the distributive law. Because this is the start it will get its own notation, we write $\rightarrowtail$, that is

$$*: U \times V \rightarrowtail W$$

denotes a distributive function on additive structures U, V, and W. As we go along we will prefer to use U, V and W in just this way so that we can get up to speed on examples as quickly as possible.

1.2 Properties of addition

It is tempting now to start assuming that U, V, and W are something family—vector spaces, modules, or at least abelian groups. However this would rob the distributive law of its power and leave us to think addition and its common attributes are the reason tensors work. But the distributive law is already claiming a strong interaction of two operations so maybe it should be explored on its own a little while longer to appreciate what it already says about the individual operations. More examples will demonstrate the value of a general point of view.

We will use a number of spaces

$$\begin{aligned} \mathbb{R}^d &:= \{u : [d] \to \mathbb{R}\}, \\ R^{m\times n} &:= \{M : [m]\times[n] \to \mathbb{R}\} \\ R^{\ell\times m\times n} &:= \{\Gamma : [\ell]\times[m]\times[n] \to \mathbb{R}\} \\ &\vdots \end{aligned}$$

Define the following operations.

$$\mathbb{R}^m \oplus \mathbb{R}^n := \mathbb{R}^{m+n}$$

$$\begin{bmatrix} \mathbb{R}^{a\times n} \\ \mathbb{R}^{b\times n} \end{bmatrix} := \mathbb{R}^{(a+b)\times n} \qquad \begin{bmatrix} \mathbb{R}^{m\times c} & \mathbb{R}^{m\times d} \end{bmatrix} := \mathbb{R}^{m\times(c+d)}$$

$$\vdots$$

Now we add a multiplication.

$$\mathbb{R}^m \otimes \mathbb{R}^n := \mathbb{R}^{m\times n}$$

Example 1.1. The distributive law with vector space operators.

$$(\mathbb{R}^a \oplus \mathbb{R}^b) \otimes \mathbb{R}^n = \begin{bmatrix} \mathbb{R}^a \otimes \mathbb{R}^n \\ \mathbb{R}^b \otimes \mathbb{R}^n \end{bmatrix}$$
$$\mathbb{R}^m \oplus (\mathbb{R}^c \oplus \mathbb{R}^d) = \begin{bmatrix} \mathbb{R}^m \otimes \mathbb{R}^c & \mathbb{R}^m \otimes \mathbb{R}^d \end{bmatrix}$$

Lets combine the left and right distributive laws.

$$\begin{array}{ccc} & (u+\acute{u})*(v+\acute{v}) & \\ (u+\acute{u})*v+(u+\acute{u})*\acute{v} & & u*(v+\acute{v})+\acute{u}*(v+\acute{v}) \\ \| & & \| \\ (u*v+\acute{u}*v)+(u*\acute{v}+\acute{u}*\acute{v}) & & (u*v+u*\acute{v})+(\acute{u}*v+\acute{u}*\acute{v}) \end{array}$$

Thus the values $a, b, c, d, \ldots$ in the image of a distributive product $*: U \times V \rightarrowtail W$ must satisfy the following identity.

$$(a+b)+(c+d) = (a+c)+(b+d) \tag{Distributable}$$

Proposition 1.2. If $+$ is both associative and commutative then it is distributable. Furthermore, such an addition can be extended to have at 0.

Conversely, if $+$ is distributable and has a 0 then $+$ is associative and commutative.

Proof. TBD □

Algebra you were told to forget

We are in a territory where we are assuming very little about our addition and we may even include structures like addition of vector spaces, which are not in a technical sense sets. This causes no trouble. Algebra predates sets (think of Boolean, Heyting and de Morgan algebra) and outlasts sets (think of categories). So what is algebra with sets? For that we go back to grammar school.

We want additive structures so we use $+$ and we want that be $x + y$ not stranger styles like $+x\ y$ or `add(x,y)` which have the place but not here. To feel sophisticated about this we could report that we have the following grammar:

```
<A>::= (<A>+<A>) | 0
```

where | indicates "or" and so we can add or use 0. If we do not want 0 in our addition we just remove it; likewise, if wish to add something like negatives we can add this as well as in the following.

```
<An>::= (<AN>+<AN>) | 0 | -<AN>
```

We will want some formulas. So we add variables, for examples to add $X = \{x, y, z\}$ as variables we use

```
   <X> ::= x | y
<A<X>> ::= <X> | 0 |(<A<X>>+<A<X>>)
```

Strings $\Phi := \Phi(X)$ that prase in this grammar are called *additive formulas* and annotated $\Phi : \mathrm{Fr}_A\langle X\rangle$. For example $((x + 0) + y) : \mathrm{Fr}_A\langle x, y\rangle$, but not $+xy$ because that does not match the pattern in the grammar.

Given an additive structure W and assignment of variables $w_X : W^X$ we can substitute into a formula $\Phi(X)$ the values w_X formula a list of operations to carry out in W called *evaluation* and denoted $\Phi(w_X)$. For example $\Phi(x, y) = ((x + 0) + y)$ and $W = \mathbb{N}$ then $\Phi(2, 3) = ((2 + 0) + 3) = 5$.

Now given an additive algebra W and a function $S : I \to W$, then the evaluations $\Phi(s_*)$ are said to have type

$$\mathrm{Fr}_{+,0}\langle S\rangle$$

We say that $\mathrm{Fr}_{+,0}\langle S\rangle$ is the subalgebra generated by S.

Distributable

Definition 1.3. Given a distributive product $* : U \times V \rightarrowtail W$ let

$$U * V = \mathrm{Fr}_{+,0}\langle u * v \mid u \in U, v \in V\rangle$$

Proposition 1.4. Given a distributive product $* : U \times V \to W$ then $U/V^\perp$, $V/U^\top$ and $W^+ = U * V$ are commonoids

Proof. TBD □