1 What is a tenser?

Let us define a tenser of rank $r \in \mathbb{N}$ to be the data:

- A set of numbers $s_1, s_2, \ldots s_r \in \mathbb{N}$, where s_i is said to be the size of the tensor along dimension i.
- A function $F : [s_1] \times [s_2] \times \cdots \times [s_r] \to (\mathbb{R} \to \mathbb{R})$, where $[u] \equiv \{1, 2, \ldots u\}$.

Given a tenser $T \equiv (r, S \equiv (s_1, s_2, \ldots, s_r), F)$, we denote by:

$$T[ix_1, ix_2, \ldots ix_r] : ([s_1] \times [s_2] \times \cdots \times [s_r]) \to (\mathbb{R} \to \mathbb{R})$$
$$T[ix_1, ix_2, \ldots ix_r] \equiv F(ix_1, ix_2, ix_3, \ldots, ix_r)$$

Let us now instantiate an honest to god tenser. We shall create what plebes know as a "vector field" to be a tenser.

1.1 Example: Vector field as a tenser

Consider the vector field on $\mathbb{R}^2$ to be $V \equiv (sin(x), cos(y))$.

For us, this will correspond to a tenser of rank 1 $T \equiv (r \equiv 1, S \equiv (2), F)$ where:

$$F(1) \equiv \lambda x. \sin x \quad F(2) \equiv \lambda x. \cos x$$

1.2 Example: Function as a tenser

A function $f : \mathbb{R} \to \mathbb{R}$ is a tenser of rank 1 $T_f \equiv (r \equiv 1, S \equiv (1), F(1) \equiv f)$.

1.3 Example: Scalar as a tenser

A scalar $r \in \mathbb{R}$ is a tenser of rank 1 $T \equiv (r \equiv 0, S \equiv (), F() \equiv \lambda\_.r)$

1.4 Matrix field as a tenser

The matrix field which maps each point (x, y) to the matrix $\begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix}$ is a rank 2 tenser:

$$T \equiv (r \equiv 2, S \equiv (2, 2), F)$$
$$F(1, 1) \equiv \lambda x.x \quad F(1, 2) \equiv \lambda\_.0$$
$$F(2, 1) \equiv \lambda\_.0 \quad F(2, 2) \equiv \lambda y.y$$

2 Tenser derivatives

We are often interested in understanding how one tenser varies with respect to another. But what does this question even mean? Well, I claim there is only one sensible explanation. A tenser after all is just a collection of functions. So the derivative of one tenser with respect to another, say $\frac{\partial A[i_1, i_2, \ldots i_n]}{\partial B[j_1, j_2, \ldots j_m]}$ can only be a new tenser whose entries are the derivatives of *each* function in A with *each* function in B.

This instantly leads to the definition:

$$
\begin{aligned}
&A \equiv (n, S_A, F_a) \quad B \equiv (m, S_B, F_b) \\
&C \equiv (n+m, (S_A, S_B), F) \\
&F(i_1, i_2, \ldots i_n, j_1, j_2, \ldots j_m) \equiv \frac{\partial A[i_1, i_2, \ldots i_n]}{\partial B[j_1, j_2, \ldots j_m]}
\end{aligned}
$$

Note that this is perfectly well defined, since $A[i_1, i_2, \ldots i_n] : \mathbb{R} \to \mathbb{R}$. Similarly, $B[j_1, j_2, \ldots j_m] : \mathbb{R} \to \mathbb{R}$, and we hopefully know how to differentiate single variable functions.

The C as written above is often colourfully written as:

$$
\begin{aligned}
&\frac{\partial A}{\partial B} \\
&\frac{\partial A[i_1, i_2, \ldots i_n]}{\partial B[j_1, j_2, \ldots j_m]}
\end{aligned}
$$

and many other abuses of notation. But never forget what it is doing: It is simply creating a convenient way to consider the change of every component of A relative to every component of B.

3 The derivative $\frac{\partial x^T x}{\partial x}$

Notice that if x is an honest to god vector, the above expression makes no sense. For example, let $x = (5, 5) \in \mathbb{R}^2$. Now, $x^T x = 50$, leading to the absurd expression $\frac{\partial 50}{\partial (5,5)}$, which is quite senseless since differentiation is only defined for *functions*.

Hence, whenever people write such expressions, they really mean a rank 1, shape (n) tenser. that is, an n-tuple of scalar functions, each function describing the value of the nth component of the vector, relative to some parametrization.

Let $x \equiv (r \equiv 1, S \equiv (2), F_x)$. Now, $x^T x \equiv (r \equiv 0, S \equiv (()), F_{xtx} \equiv F_x(1)F_x(1) + F_x(2)F_x(2) = F_x(1)^2 + F_x(2)^2)$

Let us now calculate $\frac{\partial x^T x}{\partial x}$ as we have agreed upon above:

$$
\begin{aligned}
&\frac{\partial x^T x}{\partial x}(r \equiv 0+1, S \equiv 2, F_{der}) \\
&F_{der}[][1] = \frac{\partial F_{xtx}}{\partial F_x(1)} = \frac{\partial F_x(1)^2 + F_x(2)^2}{\partial F_x(1)} = 2F_x(1) \\
&F_{der}[][2] = \frac{\partial F_{xtx}}{\partial F_x(2)} = \frac{\partial F_x(1)^2 + F_x(2)^2}{\partial F_x(2)} = 2F_x(2)
\end{aligned}
$$

Hence, $\frac{\partial x^T x}{\partial x} = 2 \cdot x$.

We can show that the collection of tensers form a vector space over $\mathbb{R}$, since functions $\mathbb{R} \to \mathbb{R}$ form a vector space over $\mathbb{R}$, and a tenser is a clever collection of such scalar functions. Hence, the notation $2 \cdot x$ is interpreted in terms of this vector space structure.