Construction of tensor product

Wu Yuhan

Sept. 28, 2020

1 Introduction

This article is aimed to present a construction of tensor product in mathematicians' way with strong motivation. Firstly, we review the property of dual space and bilinear function. Then, the tensor product will be constructed as a bilinear function. The main motivation is that we want to construct a linear function of a vector space that are combined with two or more vector spaces, and the combination is the process of construction of tensor product. Finally, the transformation of coordinates in this definition satisfy the definition of usual tensors. This article mainly refers to *Lectures on differential geometry*, written by S.S. Chern.

2 Preparation: Dual space

$\mathbf{V}$ is a vector space with dimension n, and $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ is a orthonormal basis of $\mathbf{V}$. All linear functions on $\mathbf{V}$ form a vector space , noted as $\mathbf{V}^*$. $\alpha_1^*, \alpha_2^*, \cdots, \alpha_n^*$ is the basis of $\mathbf{V}^*$ with property:

$$\alpha_i^*(\alpha_j) = \delta_{ij}$$

Let's check it, and only need to verify that $\forall f \in \mathbf{V}^*$, can be expanded by $\{\alpha_i^*\}$.

Proof *First of all,* $\forall \alpha = \sum_{j=1}^n a_j \alpha_j \in \mathbf{V}$, *note that ,*

$$\begin{aligned}
\alpha_i^*(\alpha) &= \alpha_i^*(\sum_{j=1}^n a_j \alpha_j) \\
&= \sum_{j=1}^n \alpha_i^*(\alpha_j) a_j \\
&= \sum_{j=1}^n \delta_{ij} a_j = a_i
\end{aligned}$$

Then, $\forall f \in \mathbf{V}^*$,

$$\begin{aligned}
f(\alpha) &= f(\sum_{i=1}^n \alpha_i) \\
&= \sum_{i=1}^n \alpha_i^*(\alpha) f(\alpha_i) \\
&= \sum_{i=1}^n f(\alpha_i) \alpha_i^*(\alpha)
\end{aligned}$$

Therefore, $f = \sum_{i=1}^n f(\alpha_i)\alpha_i^*$

Definition 1 *Double dual space, we can construct the dual space of* $\mathbf{V}^*$, *noted as*$\mathbf{V}^{**}$, *and* $\{\alpha_i^{**}\}$ *is the basis with similar property* $\alpha_i^{**}(\alpha_j^*) = \delta_{ij}$

Now, we want to show that the double dual space $\mathbf{V}^{**}$ can be viewed as the vector space $\mathbf{V}$ itself, because the construction of double dual space is independent of the choose of basis. This assertion is quite difficult to understand. If you have read my another article *Dual space and inner product*, we can understand this from the point of inner product. Although it is very important, reader only need to have a glimpse on it and feel free to jump to the next section.

Proposition 1 *$\forall \alpha^{**} \in \mathbf{V}^{**}$, and $\forall \beta \in \mathbf{V}^*$, we have*

$$\alpha^{**}(\beta^*) = \beta^*(\alpha)$$

Proof

$$\begin{aligned}
\alpha^{**}(\beta^*) &= \sum_{i=1}^{n} a_i \alpha_i^{**}(\sum_{j=1}^{n} b_j \alpha_j^*) \\
&= \sum_{i,j=1}^{n} a_i b_j \alpha_i^{**}(\alpha_j^*) \\
&= \sum_{i,j=1}^{n} a_i b_j \delta_{ij} \\
&= \sum_{i,j=1}^{n} a_i b_j \alpha_j^*(\alpha_i) \\
&= \sum_{j=1}^{n} b_j \alpha_j^*(\sum_{i=1}^{n} a_i \alpha_i) \\
&= \beta^*(\alpha)
\end{aligned}$$

3 Bilinear function and tensor product

First of all, we need to review the property of bilinear functions.

Definition 2 *$\alpha, \alpha_1, \alpha_2 \in \mathbf{V}$ and $\beta, \beta_1, \beta_2 \in \mathbf{W}$, bilinear function $f : \mathbf{V} \times \mathbf{W} \to \mathbf{F}$ satisfies*

$$\begin{aligned}
f(a_1\alpha_1 + a_2\alpha_2, \beta) = a_1 f(\alpha_1, beta) + a_2 f(\alpha_2, \beta) \\
f(\alpha, b_1\beta_1 + b_2\beta_2) = b_1 f(\alpha, \beta_1) + b_2 f(\alpha, \beta_2)
\end{aligned}$$

3.1 Motivation

We want to construct a new vector space $\mathbf{Y}$, and a unique linear function on $\mathbf{Y}$, such that

$$f = g \circ h \tag{1}$$

Here, bilinear mapping h is the key to construct tensor product.

3.2 Construction of tensor product

Proposition 2 *All bilinear functions form a linear space, noted as $\mathcal{L}(\mathbf{V}, \mathbf{W}; \mathbf{F})$ or $\mathcal{L}$ in short.*

Proof *We only need to check that if $f_1, f_2 \in \mathcal{L}$, then $f_1 + f_2$ and $kf_1 \in \mathcal{L}$, which means $f_1 + f_2$ and kf_1 satisfy the condition. Readers can easily finish this by definition, if you have taken the course* Mathematical proof *or the similar course. If I have enough time, I will finish the verification in appendix.*

Instead of constructing the tensor product of $\mathbf{V}$ and $\mathbf{W}$, we choose to firstly define the tensor product of the dual space$\mathbf{V}^*$ and $\mathbf{W}$.

Definition 3 *$v^* \in \mathbf{V}^*, w^* \in \mathbf{W}^*$, define the tensor product $v^* \otimes w^*$ as a bilinear function , with $\alpha \in \mathbf{V}, \beta \in \mathbf{W}$*

$$v^* \otimes w^*(\alpha, \beta) = v^*(\alpha) \cdot w^*(\beta) \tag{2}$$

Note: $\otimes$ is a map from $\mathbf{V}^* \times \mathbf{W}^*$ to $\mathcal{L}(\mathbf{L}, \mathbf{W}; \mathbf{F})$

Proposition 3 *$\otimes$ is a bilinear map*

Proof *Assume $v, v_i^* \in \mathbf{V}^*, w, w_i^* \in \mathbf{W}^*$, and $\alpha \in \mathbf{V}, \beta \in \mathbf{W}$, verify that*

$$\begin{aligned}(a_1v_1^* + a_2v_2^*) \otimes w^*(\alpha, \beta) &= (a_1v_1^* + a_2v_2^*)(\alpha) \cdot w^*(\beta) \\ &= a_1v_1^*(\alpha)w^*(\beta) + a_2v_2^*(\alpha)w^*(\beta) \\ &= a_1v_1^* \otimes w^*(\alpha, \beta) + a_2v_2^* \otimes w^*(\alpha, \beta)\end{aligned}$$

Therefore, $\otimes$ is linear in variable v^*; similarly, $\otimes$ is also linear in another variable W^*. Conclude that $\otimes$ is a bilinear map.

$\mathbf{V}^* \otimes \mathbf{W}^*$ refers to the space generated by all elements in the form of $v^* \otimes w^*$. And we want to find the basis of $\mathbf{V}^* \otimes \mathbf{W}^*$. Since $\{\alpha_i^*\}$ and $\{\beta_j^*\}$ are bases of $\mathbf{V}^*$ and $\mathbf{W}^*$, we naturally hypothesis that $\{\alpha_i^* \otimes \beta_j^*\}$ is the basis of $\mathbf{V}^* \otimes \mathbf{W}^*$ Let's check it, recall that

$$f \in \mathbf{V}^*, f = \sum_i f(\alpha_i)\alpha_i^*$$

Then,

$$\begin{aligned}v^* \otimes w^* &= \sum_i v^*(\alpha_i)\alpha_i^* \otimes w^*(\beta_j)\beta_j^* &(3)\\ &= \sum_{i,j} v^*(\alpha_i)w^*(\beta_j)\alpha_i^* \otimes \beta_j^* &(4)\end{aligned}$$

Clearly, $\mathbf{V}^* \otimes \mathbf{W}^*$ is a subspace of $\mathcal{L}(\mathbf{V}, \mathbf{W}; \mathbf{F})$, here we want to show shat

$$\mathbf{V}^* \otimes \mathbf{W}^* = \mathcal{L}(\mathbf{V}, \mathbf{W}; \mathbf{F}) \tag{5}$$

We only need to prove that $\{\alpha_i^* \otimes \beta_j^*\}$ is a basis of $\mathcal{L}(\mathbf{V}, \mathbf{W}; \mathbf{F})$.

Proposition 4 *$\forall f \in \mathcal{L}(\mathbf{V}, \mathbf{W}; \mathbf{F})$, f can be linearly expanded by $\{\alpha_i^* \otimes \beta_j^*\}$*

Proof $\forall \alpha = \sum_i a_i\alpha_i \in \mathbf{V}$, $\forall \beta = \sum_j b_j\beta_j \in \mathbf{W}$

$$\begin{aligned}f(\alpha, \beta) = f(\sum_i a_i\alpha_i, \sum_j b_j\beta_j) &= \sum_{i,j} a_ib_jf(\alpha_i, \beta_j) \\ &= \sum_{i,j} \alpha_i^*(\alpha)\beta_j^*(\beta)f(\alpha_i, \beta_j) \\ &= \sum_{i,j} f(\alpha_i, \beta_j)\alpha_i^* \otimes \beta_j^*(\alpha, \beta)\end{aligned}$$

Therefore,

$$f = \sum_{i,j} \alpha_i^* \otimes \beta_j^* \tag{6}$$

We can conclude that

$$\mathbf{V}^* \otimes \mathbf{W}^* = \mathcal{L}(\mathbf{V}, \mathbf{W}; \mathbf{F}) \tag{7}$$

3.3 Construction the tensor product :$\mathbf{V} \otimes \mathbf{W}$

Treat the dual space $\mathbf{V}^*$ and $\mathbf{W}^*$ in same procedure:

$$\mathbf{V}^{**} \otimes \mathbf{W}^{**} = \mathcal{L}(\mathbf{V}^*, \mathbf{W}^*; \mathbf{F}) \tag{8}$$

Recall that we have introduced the double dual space, we can derive that $\mathbf{V}^{**} = \mathbf{V}$, so

$$\mathbf{V} \otimes \mathbf{W} = \mathcal{L}(\mathbf{V}^*, \mathbf{W}^*; \mathbf{F}) \tag{9}$$

Let's investigate the relation between$\mathbf{V} \otimes \mathbf{W}$ and $\mathbf{V}^* \otimes \mathbf{W}^*$.

Proposition 5 *$\mathbf{V} \otimes \mathbf{W}$ and $\mathbf{V}^* \otimes \mathbf{W}^*$ are dual spaces of each other, i.e.*

$$(\mathbf{V} \otimes \mathbf{W})^* = \mathbf{V}^* \otimes \mathbf{W}^* \tag{10}$$

,by defining $v_1^* \otimes w_1^*(v \otimes w) = v_1^*(v)w_1^*(w)$, where $v \in \mathbf{V}, w \in \mathbf{W}, v_1^* \in \mathbf{V}^*, w_1^* \in \mathbf{W}^*$

Proof *First of all, by definition in the proposition, assume that $\{\alpha_i\}, \{\beta_j\}$ are bases of $\mathbf{V}, \mathbf{W}$ and $\{\alpha_i^*\}, \{\beta_j^*\}$ are bases of $\mathbf{V}, \mathbf{W}$, and we get*

$$\alpha_i^* \otimes \beta_j^*(\alpha_p \otimes \beta_q) = \alpha_i^*(\alpha_p)\beta_j^*(\beta_q) = \delta_{i,p}\delta_{j,p} \tag{11}$$

We can assert that $\{\alpha_i \otimes \beta_j\}$ and $\alpha_i^* \otimes \beta_j^*$ are dual bases.

Check $\forall f \in (\mathbf{V} \otimes \mathbf{W})^*$, note that $\mathbf{V} \otimes \mathbf{W}$ should be considered as a vector space with basis $\{\alpha_i \otimes \beta_j\}$; $\forall \alpha \otimes \beta \in \mathbf{V} \otimes \mathbf{W}$

$$\begin{aligned} f(\alpha \otimes \beta) &= f(\sum_i a_i\alpha_i \otimes \sum_j b_j\beta_j) \\ &= \sum_{i,j} a_ib_jf(\alpha_i \otimes \beta_j) = \sum_{i,j} \alpha_i^*(\alpha)\beta_j^*(\beta)f(\alpha_i \otimes \beta_j) \\ &= \sum_{i,j} f(\alpha_i \otimes \beta_j)\alpha_i^* \otimes \beta_j^*(\alpha \otimes \beta) \end{aligned}$$

So f can be linearly expanded by $\{\alpha_i^* \otimes \beta_j^*\}$, which means

$$(\mathbf{V} \otimes \mathbf{W})^* \subseteq \mathbf{V}^* \otimes \mathbf{W}^* \tag{12}$$

Similarly, consider the double dual space and apply the same procedure above, we can prove:

$$\begin{aligned} &(\mathbf{V}^* \otimes \mathbf{W}^*)^* \subseteq \mathbf{V}^{**} \otimes \mathbf{W}^{**} \\ &(\mathbf{V}^* \otimes \mathbf{W}^*)^{**} \subseteq \mathbf{V}^{**} \otimes \mathbf{W}^{**})^* \\ &\mathbf{V}^* \otimes \mathbf{W}^* \subseteq (\mathbf{V} \otimes \mathbf{W})^* \end{aligned}$$

Therefore,

$$\mathbf{V}^* \otimes \mathbf{W}^* = (\mathbf{V} \otimes \mathbf{W})^* \tag{13}$$

Let's continue to a big theorem. It is quite easy to prove, but it indicates essential parts of tensor product.

Theorem 1 *Define: $h : \mathbf{V} \times \mathbf{W} \to \mathbf{V} \otimes \mathbf{W}$ is a bilinear map, obtained from tensor product. $v \in \mathbf{V}, w \in \mathbf{W}$, then*

$$h(v, w) = v \otimes w \tag{14}$$

And, for any bilinear function $f : \mathbf{V} \times \mathbf{W} \to \mathbf{F}$, there exists only linear function $g : \mathbf{V} \otimes \mathbf{W} \to \mathbf{F}$ such that

$$f = g \circ h \tag{15}$$

Proof *First step, we need to find g. Naturally, we can define $g : \mathbf{V} \otimes \mathbf{W} \to \mathbb{F}$ by*

$$g(\alpha_i \otimes \beta_j) = f(\alpha_i, \beta_j) \tag{16}$$

Since, $\{\alpha_i \otimes \beta_j\}$ is the basis of $\mathbf{V} \otimes \mathbf{W}$, g is uniquely defined.

Next step, we need to check $f = g \circ h$, for arbitrary $\alpha = \sum_i a_i\alpha_i \in \mathbf{V}, \beta = \sum_j b_j\beta_j \in \mathbf{W}$,

$$\begin{aligned} g \circ h(\alpha, \beta) = g(\alpha \otimes \beta) &= g(\sum_i a_i\alpha_i \otimes \sum_j b_j\beta_j) \\ &= \sum_{i,j} a_ib_jg(\alpha_i \otimes \beta_j) = \sum_{i,j} a_ib_jf(\alpha_i, \beta_j) \\ &= f(\sum_i a_i\alpha_i, \sum_j b_j\beta_j) = f(\alpha, \beta) \end{aligned}$$

By tensor product, we can construct a bigger linear space $\mathbf{V} \otimes \mathbf{W}$ from $\mathbf{V}$ and $\mathbf{W}$. Also, the bilinear function $f(\alpha, \beta)$ and linear function $g(\alpha \otimes \beta)$ are connected by tensor product $\otimes$

4 Multi-linear function

In the previous section, we have shown that a tensor product of two linear function is a bilinear function. Similarly we can define the tensor product of multi-linear functions. For example, $f \in \mathcal{L}(\mathbf{V_1}, \cdots, \mathbf{V_s}; \mathbf{F})$ is a s-linear function, and $g \in \mathcal{L}(\mathbf{W_1}, \cdots, \mathbf{W_r}; \mathbf{F})$ is a r-linear function, then we can define a r+s-linear function $f \otimes g$ by

$$f \otimes g(v_1, v_2, \cdots, v_s, w_1, \cdots, w_r) = f(v_1, \cdots, v_s)g(w_1, \cdots, w_r) \tag{17}$$

We can verify $f \otimes g$ is indeed a r+s-linear function just by repeating the process in previous section, but this is quite boring, and nobody wants to spend time reading that. Hence, I will skip this, and focus on something more interesting.

Theorem 2 *Tensor product is association.* $f \in \mathcal{L}(\mathbf{V_1}, \cdots, \mathbf{V_s}; \mathbf{F})$, $g \in \mathcal{L}(\mathbf{W_1}, \cdots, \mathbf{W_r}; \mathbf{F})$ *and* $h \in \mathcal{L}(\mathbf{Z_1}, \cdots, \mathbf{Z_t}; \mathbf{F})$, *then*

$$(f \otimes g) \otimes h = f \otimes (g \otimes h) \tag{18}$$

Proof *Note that* $v = (v_1, \cdots, v_s)$, $w = (w_1, \cdots, w_r)$ *and* $z = (z_1, \cdots, z_t)$

$$\begin{aligned} (f \otimes g) \otimes h(v, w, z) &= (f \otimes g)(v, w)h(z) \\ &= (f(v)g(w))h(z) \\ &= f(v)(g(w)h(z)) \\ &= f(v)(g \otimes h)(v, z) \\ &= f \otimes (g \otimes h)(v, w, z) \end{aligned}$$

Therefore, $\mathbf{V_1}, \mathbf{V_2}, \cdots, \mathbf{V_s}$ are vector spaces, and we can define their tensor product and construct a bigger vector space

$$\mathbf{V_1} \otimes \mathbf{V_2} \cdots \otimes \mathbf{V_s}$$

with basis $\{\alpha^{(1)}_{j_1} \otimes \alpha^{(2)}_{j_2} \otimes \cdots \otimes \alpha^{(s)}_{j_s}\}$, where $\{\alpha^{(t)}_{j_t}\}$ is the basis of $\mathbf{V}_t$

5 Tensors

Finally, we are approaching to our ultimate goal, tensors. In this section, tensors are naturally defined when changing the basis of vector space and their tensor product. To make problems simple, and make this interesting, we only study the tensor product of $\mathbf{V}$ and its dual space $\mathbf{V}^*$

Suppose that $\mathbf{V}$ is a n-dimensional vector space and $\mathbf{V}^*$ is its dual space.

Definition 4 *(r,s)-type tensors:*

$$\mathbf{V}^r_s = \otimes^r_{i=1}\mathbf{V} \otimes^s_{j=1} \mathbf{V}^* = \overbrace{\mathbf{V} \otimes \cdots \otimes \mathbf{V}}^{r\ terms} \otimes \overbrace{\mathbf{V}^* \otimes \cdots \otimes \mathbf{V}^*}^{s\ terms} \tag{19}$$

where r is contravariant order and s is covariant order.

Here are some examples:$\mathbf{V}^0_0 = \mathbf{F}$, $\mathbf{V}^1_0 = \mathbf{V}$ and $\mathbf{V}^0_1 = \mathbf{V}^*$

Let's recall that $\mathbf{V} \otimes \mathbf{W} = \mathcal{L}(\mathbf{V}^*, \mathbf{W}^*; F)$

The elements in V^r_s can be viewed as a (r+s)-linear function defined on

$$\overbrace{\mathbf{V}^* \times \cdots \mathbf{V}^*}^{r} \times \overbrace{\mathbf{V} \times \cdots \times \mathbf{V}}^{s}$$

In other words,

$$\mathbf{V}^r_s = \mathcal{L}(\overbrace{\mathbf{V}^* \times \cdots \mathbf{V}^*}^{r} \times \overbrace{\mathbf{V} \times \cdots \times \mathbf{V}}^{s}; \mathbf{F}) \tag{20}$$

$\mathbf{V}^r_s$ has a basis $e_{i_1} \otimes e_{i_2} \cdots \otimes e_{i_r} \otimes e^{*k_1} \otimes \cdots \otimes e^{*k_s}$, where $1 \leqslant i_1, \cdots, i_r, k_1, \cdots, k_s \leqslant n$

Recall the first big theorem, we can view the tensor product as a linear function on its dual space. For example, $v \otimes w$ is a linear function on $\mathbf{V}^* \otimes \mathbf{W}^*$.

So the element in $\mathbf{V}_s^r$ can be viewed as a linear function on vector space

$$\overbrace{\mathbf{V}^* \otimes \cdots \otimes \mathbf{V}^*}^{r} \otimes \overbrace{\mathbf{V} \otimes \cdots \otimes \mathbf{V}}^{s}$$

Let's have a closer view

$$e_{i_1} \otimes e_{i_2} \cdots \otimes e_{i_r} \otimes e^{*k_1} \otimes \cdots \otimes e^{*k_s}(e^{*j_1} \otimes \cdots \otimes e^{*j_r} \otimes e_{m_1 \otimes \cdots \otimes e_{m_s}}) = \delta_{i_1,j_1} \cdots \delta_{i_r,j_r} \delta_{k_1,m_1} \cdots \delta_{k_s,m_s} \tag{21}$$

$\forall x \in \mathbf{V}_s^r$

$$x = \sum_{i_1,\cdots,i_r,k_1,\cdots k_s} x^{i_1,\cdots,i_r}_{k_1,\cdots,k_s} e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{*k_1} \otimes \cdots \otimes e^{*k_s}$$
$$\text{where } x^{i_1,\cdots,i_r}_{k_1,\cdots,k_s} = x(e^{*i_1} \otimes \cdots \otimes e^{*i_r} \otimes e_{k_1} \otimes \cdots \otimes e_{k_s})$$

Change of basis and the definition of tensors

First of all, consider the change of basis

$$\bar{e}_i = \sum_l a_i^l e_l \tag{22}$$

$$\bar{e}^{*j} = \sum_k b_k^j e^{*k} \tag{23}$$

Proposition 6 *We have the relation :*

$$\delta_{ij} = \sum_l b_l^j a_i^l \tag{24}$$

Proof

$$\begin{aligned}\delta_{ij} = \bar{e}^{*j}(\bar{e}_i) &= \sum_k b_k^j e^{*k}(\sum_l a_i^l e_l) \\ &= \sum_{k,l} b_k^j a_i^l e^{*k}(e_i) \\ &= \sum_{k,l} b_k^j a_i^l \delta_{k,l} \\ &= \sum_l b_l^j a_i^l\end{aligned}$$

Expand $x \in \mathbf{V}_s^r$ on bases $\{\bar{e}_{i_1} \otimes \cdots \otimes \bar{e}_{i_r} \otimes \bar{e}^{*k_1} \otimes \cdots \otimes \bar{e}^{*k_s}\}$and $\{e_{j_i} \otimes \cdots \otimes e_{j_r} \otimes e^{*l_1} \otimes \cdots \otimes e^{l_s}\}$ respectively.

$$x = \sum_{\substack{i_1,\cdots i_r \\ k_1 \cdots k_s}} \bar{x}^{i_1,\cdots,i_r}_{k_1,\cdots,k_s} \bar{e}_{i_1} \otimes \cdots \otimes \bar{e}_{i_r} \otimes \bar{e}^{*k_1} \otimes \cdots \otimes \bar{e}^{*k_s} \tag{25}$$

$$= \sum_{\substack{j_1 \cdots j_r \\ l_1 \cdots l_s}} \sum_{\substack{i_1,\cdots i_r \\ k_1 \cdots k_s}} \bar{x}^{i_1,\cdots,i_r}_{k_1,\cdots,k_s} a^{j_1}_{i_1} \cdots a^{j_r}_{i_r} b^{k_1}_{l_1} \cdots b^{k_s}_{l_s} e_{j_i} \otimes \cdots \otimes e_{j_r} \otimes e^{*l_1} \otimes \cdots \otimes e^{l_s} \tag{26}$$

$$= \sum_{\substack{j_1 \cdots j_r \\ l_1 \cdots l_s}=1}^{n} x^{j_1 \cdots j_r}_{k_1 \cdots k_s} e_{j_i} \otimes \cdots \otimes e_{j_r} \otimes e^{*l_1} \otimes \cdots \otimes e^{l_s} \tag{27}$$

So we can get

$$x^{j_1 \cdots j_r}_{k_1 \cdots k_s} = \sum_{\substack{i_1,\cdots i_r \\ k_1 \cdots k_s}} \bar{x}^{i_1,\cdots,i_r}_{k_1,\cdots,k_s} a^{j_1}_{i_1} \cdots a^{j_r}_{i_r} b^{k_1}_{l_1} \cdots b^{k_s}_{l_s} \tag{28}$$

and this is the classical definition of tensors.