HW 9 Tensor Calculations

Liu Mingdao 2020011156

Exercise 1.9.1

1. For arbitrary H_A, H_B, under the inner product structure on H_A, H_B, we pick an orthonormal basis: $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_m\}$ for H_A, $\{\tilde{\boldsymbol{e}}_1, \cdots, \tilde{\boldsymbol{e}}_n\}$ for H_B. Then for $\forall M, N \in H_A \otimes H_B$, suppose $M = \sum_{i,j} m_{ij}\boldsymbol{e}_i \otimes \tilde{\boldsymbol{e}}_j,\ N = \sum_{k,l} n_{kl}\boldsymbol{e}_k \otimes \tilde{\boldsymbol{e}}_l$. Note that under this orthonormal basis $(\boldsymbol{e}_i, \boldsymbol{e}_k) = \delta_{ik},\ (\tilde{\boldsymbol{e}}_j, \tilde{\boldsymbol{e}}_l) = \delta_{jl}$, we have

$$
\begin{aligned}
(M, N) &= (\sum_{i,j} m_{ij}\boldsymbol{e}_i \otimes \tilde{\boldsymbol{e}}_j, \sum_{k,l} n_{kl}\boldsymbol{e}_k \otimes \tilde{\boldsymbol{e}}_l) \\
&= \sum_{i,j}\sum_{k,l} m_{ij}n_{kl}(\boldsymbol{e}_i, \boldsymbol{e}_k)(\tilde{\boldsymbol{e}}_j, \tilde{\boldsymbol{e}}_l) \\
&= \sum_{i,j} m_{ij}n_{ij} \\
(N, M) &= (\sum_{k,l} n_{kl}\boldsymbol{e}_k \otimes \tilde{\boldsymbol{e}}_l, \sum_{i,j} m_{ij}\boldsymbol{e}_i \otimes \tilde{\boldsymbol{e}}_j) \\
&= \sum_{i,j}\sum_{k,l} m_{ij}n_{kl}(\boldsymbol{e}_k, \boldsymbol{e}_i)(\tilde{\boldsymbol{e}}_l, \tilde{\boldsymbol{e}}_j) \\
&= \sum_{i,j} m_{ij}n_{ij}
\end{aligned}
$$

Thus $(M, N) = (N, M)$, the structure is symetric.

$$
\begin{aligned}
(M, M) &= (\sum_{i,j} m_{ij}\boldsymbol{e}_i \otimes \tilde{\boldsymbol{e}}_j, \sum_{k,l} m_{kl}\boldsymbol{e}_k \otimes \tilde{\boldsymbol{e}}_l) \\
&= \sum_{i,j}\sum_{k,l} m_{ij}m_{kl}(\boldsymbol{e}_i, \boldsymbol{e}_k)(\tilde{\boldsymbol{e}}_j, \tilde{\boldsymbol{e}}_l) \\
&= \sum_{i,j} m_{ij}^2
\end{aligned}
$$

Thus $(M, M) \geq 0$, and $(M, M) = 0 \iff (\forall i, j) m_{ij} = 0 \iff M = \mathbf{0}$. The structure is positive definite.

2. Let $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ be the orthonormal basis in $\mathbb{R}^2$, $\boldsymbol{e}_{ij} = \boldsymbol{e}_i \otimes \boldsymbol{e}_j$. Suppose a, b are both non-zero and $rank(a\boldsymbol{e}_{11} + b\boldsymbol{e}_{22}) = 1$. For some $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^2$, $a\boldsymbol{e}_{11} + b\boldsymbol{e}_{22} = \boldsymbol{v} \otimes \boldsymbol{w} = v_1w_1\boldsymbol{e}_{11} + v_1w_2\boldsymbol{e}_{12} + v_2w_1\boldsymbol{e}_{21} + v_2w_2\boldsymbol{e}_{22}$. Compare the coefficients and arrange them in a matrix, we have $\boldsymbol{v}\boldsymbol{w}^T = diag(a, b)$ while $rank(\boldsymbol{v}\boldsymbol{w}^T) = 1, rank(diag(a, b)) = 2$, contradiction. So when $a \neq 0,\ b \neq 0$, $rank(a\boldsymbol{e}_{11} + b\boldsymbol{e}_{22}) = 2$.

3. When $rank(\omega) = 1$, $(\omega, L \otimes I_B(\omega)) = (L\boldsymbol{v} \otimes \boldsymbol{w}, \boldsymbol{v} \otimes \boldsymbol{w}) = (v_1^2 - v_2^2)(w_1^2 + w_2^2)$, $(\omega, I_A \otimes L(\omega)) = (\boldsymbol{v} \otimes \boldsymbol{w}, \boldsymbol{v} \otimes L\boldsymbol{w}) = (v_1^2 + v_2^2)(w_1^2 - w_2^2)$.

Then we prove that $(v_1^2 - v_2^2)(w_1^2 + w_2^2)$ and $(v_1^2 - v_2^2)(w_1^2 + w_2^2)$ can be any pair of real numbers. Suppose $\forall x, y \in \mathbb{R}$, let $R = \sqrt{x^2 + y^2}$. We add two constraint

$$
\begin{cases}
||\boldsymbol{v}|| = v_1^2 + v_2^2 = R,\ v_1 = Rcos\alpha,\ v_2 = Rsin\alpha \\
||\boldsymbol{w}|| = w_1^2 + w_2^2 = R,\ w_1 = Rcos\beta,\ w_2 = Rsin\beta
\end{cases}
$$

Then let

$$\begin{cases} x = (\omega, I_A \otimes L(\omega)) \\ y = (\omega, L \otimes I_B(\omega)) \end{cases}$$

We have

$$\begin{cases} \frac{x}{R} = cos2\beta \\ \frac{y}{R} = cos2\alpha \end{cases}$$

I.e. in $[0, \pi)$

$$\begin{cases} \beta = \frac{1}{2} arccos \frac{x}{R} \\ \alpha = \frac{1}{2} arccos \frac{y}{R} \end{cases}$$

Then we get $\boldsymbol{v}, \boldsymbol{w}$. Thus, for any given pair of real numbers, we can find $\boldsymbol{v}, \boldsymbol{w}$.

4. $(\omega, L \otimes I_B(\omega)) = (a\boldsymbol{e}_{11} + b\boldsymbol{e}_{22}, L \otimes I_B(a\boldsymbol{e}_{11}) + L \otimes I_B(b\boldsymbol{e}_{22})) = (a\boldsymbol{e}_{11} + b\boldsymbol{e}_{22}, a\boldsymbol{e}_{11} - b\boldsymbol{e}_{22}) = a^2 - b^2$. $(\omega, I_A \otimes L(\omega)) = (a\boldsymbol{e}_{11} + b\boldsymbol{e}_{22}, I_A \otimes L(a\boldsymbol{e}_{11}) + I_A \otimes L(b\boldsymbol{e}_{22})) = (a\boldsymbol{e}_{11} + b\boldsymbol{e}_{22}, a\boldsymbol{e}_{11} - b\boldsymbol{e}_{22}) = a^2 - b^2$. □

Exercise 1.9.2

1. Let the old basis $\mathcal{B}$ be $[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n]$, the new basis $\mathcal{C}$ be $[\boldsymbol{w}_1, \cdots, \boldsymbol{w}_n]$, then from $\boldsymbol{v}_\mathcal{C} = B\boldsymbol{v}_\mathcal{B}$, we have $[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n] = [\boldsymbol{w}_1, \cdots, \boldsymbol{w}_n]B$. Let $B^{-1} = [\hat{b_1}, \ \cdots, \ \hat{b_n}]$. Then $\boldsymbol{w}_i = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n]\hat{b_i}$, which means $\boldsymbol{w}_i = \sum\limits_{k=1}^{n} (\hat{b_i})_k \boldsymbol{v}_k$.

Under the old basis $\alpha_\mathcal{B} = [\alpha(\boldsymbol{v}_1), \ \cdots, \ \alpha(\boldsymbol{v}_n)]$. Under the new basis

$$\begin{aligned} \alpha_\mathcal{C} &= [\alpha(\boldsymbol{w}_1), \ \cdots, \ \alpha(\boldsymbol{w}_n)] \\ &= [\alpha([\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n]\hat{b_1}), \ \cdots, \ \alpha([\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n]\hat{b_n})] \\ &= [\alpha([\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n])\hat{b_1}, \ \cdots, \ \alpha([\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n])\hat{b_n}] \\ &= \alpha_\mathcal{B}[\hat{b_1}, \ \cdots, \ \hat{b_n}] \\ &= \alpha_\mathcal{B} B^{-1} \end{aligned}$$

2. Note that $\boldsymbol{v}_\mathcal{C} = B\boldsymbol{v}_\mathcal{B}$, $\boldsymbol{w}_\mathcal{C} = B\boldsymbol{w}_\mathcal{B}$. We have $L(\boldsymbol{v}_\mathcal{B} \otimes \boldsymbol{w}_\mathcal{B}) = (B\boldsymbol{v}_\mathcal{B}) \otimes (B\boldsymbol{w}_\mathcal{B})$ for $\forall \boldsymbol{v}_\mathcal{B}, \boldsymbol{w}_\mathcal{B} \in \mathbb{R}^n$. So L is defined. To determine linear map L clearly, we just need to find its corresponding matrix. Let $B = [b_1, \ \cdots, \ b_n]$. Note that $L(\boldsymbol{e}_i \otimes \boldsymbol{e}_j) = (B\boldsymbol{e}_i) \otimes (B\boldsymbol{e}_j) = b_i \otimes b_j$, then the matrix of L is

$$[\ \underbrace{b_1 \otimes b_1, b_1 \otimes b_2, \cdots, b_1 \otimes b_n}_{n \text{ vectors as a group}}, \cdots, \ \underbrace{b_n \otimes b_1, b_n \otimes b_2, \cdots, b_n \otimes b_n}_{n \text{ vectors as a group}}\ ]_{n^2 \times n^2}$$

3. In the previous subproblems, we note that during the change of basis with matrix B, the coordinate of vectors in V is multiplied by B from the left, and the coordinate of dual vectors in V^* is multiplied by B^{-1} from the right. Then we infer that $(\boldsymbol{v}_1)_\mathcal{B} \otimes \cdots \otimes (\boldsymbol{v}_a)_\mathcal{B} \otimes (\alpha_1)_\mathcal{B} \otimes \cdots \otimes (\alpha_b)_\mathcal{B}$ is sent to

$$B(\boldsymbol{v}_1)_\mathcal{B} \otimes \cdots \otimes B(\boldsymbol{v}_a)_\mathcal{B} \otimes (\alpha^1)_\mathcal{B} B^{-1} \otimes \cdots \otimes (\alpha^b)_\mathcal{B} B^{-1}$$

□

Exercise 1.9.3 Just prove the generic case. Consider differentiable function $f : \mathbb{R}^3 \to \mathbb{R}$. Under the new coordinate system, we assume

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = B \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

I.e.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = B^{-1} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Take total derivative for x, y, z with a, b, c as independent variables, then arrange them in a matrix, we get

$$\begin{bmatrix} \frac{\partial x}{\partial a} & \frac{\partial x}{\partial b} & \frac{\partial x}{\partial c} \\ \frac{\partial y}{\partial a} & \frac{\partial y}{\partial b} & \frac{\partial y}{\partial c} \\ \frac{\partial z}{\partial a} & \frac{\partial z}{\partial b} & \frac{\partial z}{\partial c} \end{bmatrix} = B^{-1} \tag{1}$$

From the chain rule of partial derivative, we have

$$\begin{aligned} \frac{\partial f}{\partial a} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial a} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial a} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial a} \\ \frac{\partial f}{\partial b} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial b} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial b} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial b} \\ \frac{\partial f}{\partial c} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial c} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial c} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial c} \end{aligned}$$

I.e.

$$\begin{bmatrix} \frac{\partial f}{\partial a} \\ \frac{\partial f}{\partial b} \\ \frac{\partial f}{\partial c} \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial a} & \frac{\partial y}{\partial a} & \frac{\partial z}{\partial a} \\ \frac{\partial x}{\partial b} & \frac{\partial y}{\partial b} & \frac{\partial z}{\partial b} \\ \frac{\partial x}{\partial c} & \frac{\partial y}{\partial c} & \frac{\partial z}{\partial c} \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{bmatrix} \tag{2}$$

Compare (1) and (2), we get

$$\nabla f_{new}(a, b, c) = \begin{bmatrix} \frac{\partial f}{\partial a} \\ \frac{\partial f}{\partial b} \\ \frac{\partial f}{\partial c} \end{bmatrix} = (B^{-1})^T \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{bmatrix} = (B^{-1})^T(\nabla f_{old}(x, y, z))$$

Hence, if we arrange the gradient as row vectors, we just take transpose on both sides:

$$\begin{bmatrix} \frac{\partial f}{\partial a} & \frac{\partial f}{\partial b} & \frac{\partial f}{\partial c} \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{bmatrix} B^{-1}$$

Clearly the coordinate of gradient bahaves like the dual vectors. $\square$