2022 春高等代数习题课

2022-6-17 第七次习题课

第一题. 习题 9.5 第 9 题. 实数域 $\mathbb{R}$ 上全体 n 阶方阵构成 $\mathbb{R}$ 上 n^2 维线性空间 $V=\mathbb{R}^{n\times n}$, 在 V 中定义了内积 $\langle X, Y\rangle=\operatorname{tr}\left(XY^T\right)$ 之后称为欧氏空间。对 V 上如下的线性函数 f, 求 $B\in V$ 使 $f(X)=\langle X, B\rangle$.

(1) $f(X)=\operatorname{tr} X$;

(2) 对给定的 $A, D\in\mathbb{R}^{n\times n}$, $f(X)=\operatorname{tr}(AXD)$;

(3) 对给定的 $A, D\in\mathbb{R}^{n\times n}$, $f(X)=\operatorname{tr}(AX-XD)$;

(4) 对给定的 $\alpha, \beta\in\mathbb{R}^{1\times n}$, $f(X)=\alpha X\beta^T$.

解. (1). 要满足 $\operatorname{tr}\left(XB^T\right)=\langle X, B\rangle=f(X)=\operatorname{tr} X$, 只要令 $B=I$ 即可。

(2). 要满足 $\operatorname{tr}\left(XB^T\right)=\langle X, B\rangle=f(X)=\operatorname{tr}(AXD)=\operatorname{tr}(XDA)$, 只要令 $B^T=DA$, 即 $B=(DA)^T=A^TD^T$ 即可。

(3). 要满足

$$\begin{aligned}\operatorname{tr}\left(XB^T\right)&=\langle X, B\rangle=f(X)=\operatorname{tr}(AX-XD)\\&=\operatorname{tr}(AX)-\operatorname{tr}(XD)=\operatorname{tr}(XA)-\operatorname{tr}(XD)\\&=\operatorname{tr}(X(A-D))\end{aligned}$$

只要令 $B^T=A-D$, 即 $B=(A-D)^T$ 即可。

(4). 要满足 $\operatorname{tr}\left(XB^T\right)=\langle X, B\rangle=f(X)=\operatorname{tr}\left(\alpha X\beta^T\right)=\operatorname{tr}\left(X\beta^T\alpha\right)$, 只要令 $B^T=\beta^T\alpha$, 即 $B=(\beta^T\alpha)^T=\alpha^T\beta$ 即可。

这题主要利用了方阵迹的性质 $\operatorname{tr}(AB)=\operatorname{tr}(BA)$. 注意这里的 A, B 并不要求都是方阵，而只要求相应的矩阵乘法有定义即可，即 A 为 $m\times n$ 阶矩阵，B 为 $n\times m$ 阶矩阵即可。

第二题. 习题 9.7 第 8 题. 设 U 为酉方阵，且 $U^{-1}AU=B$, 证明：$\operatorname{tr}(A^*A)=\operatorname{tr}(B^*B)$.

证明：因为 U 为酉方阵，所以 $U^{-1}=U^*$, 那么

$$\begin{aligned}\operatorname{tr}(B^*B)&=\operatorname{tr}\left(\left(U^{-1}AU\right)^*B\right)=\operatorname{tr}\left((U^*AU)^*B\right)\\&=\operatorname{tr}(U^*A^*U^{**}B)=\operatorname{tr}(U^*A^*UB)\\&=\operatorname{tr}(A^*UBU^*)=\operatorname{tr}(A^*U(U^*AU)U^*)\\&=\operatorname{tr}(A^*(UU^*)A(UU^*))\\&=\operatorname{tr}(A^*A).\end{aligned}$$

这里，我们又用到了方阵迹的性质 $\mathrm{tr}(AB)=\mathrm{tr}(BA)$.

第三题. 习题 9.7 第 9 题. 设 H_1, H_2 都是 n 阶正定 Hermite 方阵，且 H_1-H_2 正定，求证：$H_2^{-1}-H_1^{-1}$ 正定。

证明：由于 H_1 是 n 阶正定 Hermite 方阵，那么它共轭相合于 n 阶单位阵，即存在可逆的 n 阶复方阵 P, 使得 $P^*H_1P=I_n$. 令 $\widetilde{H}_2=P^*H_2P$, 由于 H_2 是正定 Hermite 方阵，那么

- $\widetilde{H}_2^*=(P^*H_2P)^*=P^*H_2^*P=P^*H_2P=\widetilde{H}_2^*$;
- 任取非零 $x\in\mathbb{C}^n$, $\langle x, \widetilde{H}_2x\rangle=\langle x, P^*H_2Px\rangle=\langle Px, H_2Px\rangle>0$.

所以 $\widetilde{H}_2$ 也是正定 Hermite 方阵，酉相似于对角线全为正实数的对角阵，即存在酉方阵 U, 使得

$$U^*\widetilde{H}_2U=\mathrm{diag}(\lambda_1,\ldots,\lambda_n),$$

$\lambda_1,\ldots,\lambda_n>0$ 为 $\widetilde{H}_2$ 的特征值。

令 $Q=PU$, 那么

$$\begin{aligned}Q^*H_1Q&=U^*\left(P^*H_1P\right)U=U^*I_nU=U^*U=I_n,\\ Q^*H_2Q&=U^*\left(P^*H_2P\right)U=U^*\widetilde{H}_2U=\mathrm{diag}(\lambda_1,\ldots,\lambda_n).\end{aligned}$$

从而有

$$Q^*(H_1-H_2)Q=I_n-\mathrm{diag}(\lambda_1,\ldots,\lambda_n)=\mathrm{diag}(1-\lambda_1,\ldots,1-\lambda_n).$$

因为 H_1-H_2 正定，所以 $Q^*(H_1-H_2)Q$ 也是正定的，所以

$$0<\lambda_1,\ldots,\lambda_n<1.$$

那么有

$$\begin{aligned}Q^{-1}\left(H_2^{-1}-H_1^{-1}\right)(Q^*)^{-1}&=Q^{-1}H_2^{-1}(Q^*)^{-1}-Q^{-1}H_1^{-1}(Q^*)^{-1}\\ &=(Q^*H_2Q)^{-1}-(Q^*H_1Q)^{-1}\\ &=(\mathrm{diag}(\lambda_1,\ldots,\lambda_n))^{-1}-(I_n)^{-1}\\ &=\mathrm{diag}(\lambda_1^{-1}-1,\ldots,\lambda_n^{-1}-1)\end{aligned}$$

也是正定阵。

第四题. 习题 9.8 第 7 题. 设数域 $\mathbb{F}$ 上 n 维线性空间 V 上定义了非退化斜对称双线性函数 f.

(1) 证明：n 是偶数。

(2) 证明：f 在 V 的适当的基 M 下的矩阵是 $H=\begin{pmatrix}O & I_{(m)}\\ -I_{(m)} & O\end{pmatrix}$, 其中 $m=\dfrac{n}{2}$.

(3) 证明：$\mathcal{A}$ 是 V 上的辛变换 $\Leftrightarrow$ $\mathcal{A}$ 在基 M 下的矩阵 A 满足条件 $A^THA=H$. 这里辛空间 (V,f) 上的一个辛变换 $\mathcal{A}$ 换指的是保这种辛结构的线性变换，即

$$f(\mathcal{A}\alpha,\mathcal{A}\beta)=f(\alpha,\beta),\quad\forall\alpha,\beta\in V.$$

(4) 满足条件 $A^THA=H$ 的方阵称为辛方阵。要使以下方阵:

$$\begin{pmatrix} P & O \\ O & Q \end{pmatrix}, \quad \begin{pmatrix} I & X \\ O & I \end{pmatrix}, \quad \begin{pmatrix} I & O \\ X & I \end{pmatrix}$$

是辛方阵，其中的 m 阶块 P,Q,X 应当满足什么样的充分必要条件?

证明. (1) f 是斜对称双线性函数，那么任取 $u,v\in\mathbb{F}^n$, $f(u,v)=-f(v,u)$. 任取 $\mathbb{F}^n$ 的一组基 $\varepsilon_1,\dots,\varepsilon_n$, 令

$$\begin{aligned} u &= a_1\varepsilon_1+\cdots+a_n\varepsilon_n, \\ v &= b_1\varepsilon_1+\cdots+b_n\varepsilon_n, \end{aligned}$$

那么

$$\begin{aligned} f(u,v) &= f(a_1\varepsilon_1+\cdots+a_n\varepsilon_n, b_1\varepsilon_1+\cdots+b_n\varepsilon_n) \\ &= \sum_{1\leqslant i,j\leqslant n} a_ib_jf(\varepsilon_i,\varepsilon_j)=\alpha^TH\beta \end{aligned}$$

其中 $\alpha=(a_1,\dots,a_n)^T,\beta=(b_1,\cdots,b_n)^T,H=\begin{pmatrix} f(\varepsilon_1,\varepsilon_1) & \cdots & f(\varepsilon_1,\varepsilon_n) \\ \vdots & & \vdots \\ f(\varepsilon_n,\varepsilon_1) & \cdots & f(\varepsilon_n,\varepsilon_n) \end{pmatrix}$. 由于 $f(\varepsilon_i,\varepsilon_j)=-f(\varepsilon_j,\varepsilon_i)$, 所以 H 是反对称阵。

f 是非退化的，也就是说任取非零的 $u\in\mathbb{F}^n$, 都存在 $v\in\mathbb{F}^n$, 使得 $f(u,v)\neq 0$. 假设 $\det H=0$, 那么 H 列不满秩，即存在不全为 0 的 $t_1,\dots,t_n$, 使得

$$\begin{aligned} 0 &= t_1\begin{pmatrix} f(\varepsilon_1,\varepsilon_1) \\ \vdots \\ f(\varepsilon_n,\varepsilon_1) \end{pmatrix}+\cdots+t_n\begin{pmatrix} f(\varepsilon_1,\varepsilon_n) \\ \vdots \\ f(\varepsilon_n,\varepsilon_n) \end{pmatrix} \\ &= \begin{pmatrix} f(\varepsilon_1,t_1\varepsilon_1+\cdots+t_n\varepsilon_n) \\ \vdots \\ f(\varepsilon_n,t_1\varepsilon_1+\cdots+t_n\varepsilon_n) \end{pmatrix}, \end{aligned}$$

那么对非零向量 $u=t_1\varepsilon_1+\cdots+t_n\varepsilon_n$, 它与基中所有向量在双线性型 f 的作用下都等于 0, 故不存在向量 v 使得 $f(u,v)\neq 0$, 这与 f 非退化矛盾。

综上，H 是一个可逆的反对称阵，于是

$$\det H=\det H^T=\det(-H)=(-1)^n\det H,$$

故 n 为偶数。

(2) 根据 (1) 小问，要证明 f 在 V 的适当的基 $M=\{\varepsilon_1,\dots,\varepsilon_n\}$ 下的矩阵是 $H=\begin{pmatrix} O & I_{(m)} \\ -I_{(m)} & O \end{pmatrix}=\begin{pmatrix} f(\varepsilon_1,\varepsilon_1) & \cdots & f(\varepsilon_1,\varepsilon_n) \\ \vdots & & \vdots \\ f(\varepsilon_n,\varepsilon_1) & \cdots & f(\varepsilon_n,\varepsilon_n) \end{pmatrix}$, 只要证明

$$f(\varepsilon_i,\varepsilon_j)=\begin{cases} 1, & j=i+m, \\ -1, & j=i-m\ \text{(由上一条件，自动满足)}, \\ 0, & \text{其余情况} \end{cases}$$

我们任取一个非零向量 α_1, 由于 f 非退化，则存在非零向量 α_2, 使得 $f(\alpha_1,\alpha_2)=a\neq 0$. 那么我们可以考虑令

$$\varepsilon_1=\alpha_1,\quad \varepsilon_{1+m}=\frac{1}{a}\varepsilon_2,$$

即可以满足 $f(\varepsilon_1,\varepsilon_{1+m})=1$. 考虑 $\varepsilon_1,\varepsilon_{1+m}$ 生成的子空间 $W=\mathrm{span}\{\varepsilon_1,\varepsilon_{1+m}\}$. 令

$$\begin{aligned}W'&=\{\alpha\in V\mid f(\alpha,\varepsilon_1)=f(\alpha,\varepsilon_{1+m})=0\}\\&=\ker(f(\varepsilon_1,\cdot))\cap\ker(f(\varepsilon_{1+m},\cdot)).\end{aligned}$$

对 V 的任一组基 M_0, 以及 f 在这组基下的矩阵 H_0, 在 (1) 小问中已经证明了若 f 非退化，则 H_0 可逆。对 $\alpha\in V$, 令 α_{M_0} 为 α 在这组基下的坐标表示。那么

$$\ker(f(\alpha,\cdot))\cong\left\{x\in\mathbb{F}^n \mid \left(\alpha_{M_0}^T H_0\right)x=0\right\}$$

维数等于 $n-1$. 从而知 $\dim W'=n-2=2(m-1)$. 注意，此时也会有 $V=W\oplus W'$.

于是，可以考虑对 m 进行归纳证明。$m=1$ 时，可以类似 W 进行证明。若对所有的 $n-2=2(m-1)$ 维空间都证明了题干中的结论，那么存在 W' 的一组基 $M'=\{\varepsilon_1',\dots,\varepsilon_{n-2}'\}$ 使得 f' 在这组基下矩阵是 $H'=\begin{pmatrix}O & I_{(m-1)}\\ -I_{(m-1)} & O\end{pmatrix}=\begin{pmatrix}f(\varepsilon_1',\varepsilon_1') & \cdots & f(\varepsilon_1',\varepsilon_{n-2}')\\ \vdots & & \vdots\\ f(\varepsilon_{n-2}',\varepsilon_1') & \cdots & f(\varepsilon_{n-2}',\varepsilon_{n-2}')\end{pmatrix}$, 其中 f' 为 f 在 W' 上的限制。那么令

$$M=\left\{\varepsilon_1,\varepsilon_1',\dots,\varepsilon_{m-1}',\varepsilon_{m+1},\varepsilon_m',\dots,\varepsilon_{n-2}'\right\},$$

f 在这组基下的矩阵即为 $H=\begin{pmatrix}O & I_{(m)}\\ -I_{(m)} & O\end{pmatrix}$.

(3) 辛空间 (V,f) 上的一个辛变换 $\mathcal{A}$ 换指的是保这种辛结构的线性变换，即

$$f(\mathcal{A}\alpha,\mathcal{A}\beta)=f(\alpha,\beta),\quad \forall\alpha,\beta\in V.$$

对 $\alpha\in V$, 令 α_M 为 α 在这组基下的坐标表示。

$\Longrightarrow$: 假设 $\mathcal{A}$ 是 (V,f) 上的辛变换，那么任取 $\alpha,\beta\in V$, 有

$$\alpha_M^T A^T HA\beta_M=(A\alpha_M)^T H(A\beta_M)=f(\mathcal{A}\alpha,\mathcal{A}\beta)=f(\alpha,\beta)=\alpha_M^T H\beta_M$$

因为 α,β 是任取的，所以有 $A^THA=H$.

$\Longleftarrow$: 若 $A^THA=H$, 那么任取 $\alpha,\beta\in V$, 有

$$f(\mathcal{A}\alpha,\mathcal{A}\beta)=(A\alpha_M)^TH(A\beta_M)=\alpha_M^TA^THA\beta_M=\alpha_M^TH\beta_M=f(\alpha,\beta).$$

因为 α,β 是任取的，所以 $\mathcal{A}$ 是辛变换。

(4) 要使得 $\begin{pmatrix}P & O\\ O & Q\end{pmatrix}$ 是辛方阵，依定义，需要满足

$$\begin{pmatrix}O & I_{(m)}\\ -I_{(m)} & O\end{pmatrix}=H=\begin{pmatrix}P & O\\ O & Q\end{pmatrix}^T H\begin{pmatrix}P & O\\ O & Q\end{pmatrix}=\begin{pmatrix}O & P^TQ\\ -Q^TP & O\end{pmatrix},$$

即满足 $P^TQ=I_{(m)}$ 即可。

要使得 $\begin{pmatrix}I & X\\ O & I\end{pmatrix}$ 是辛方阵，依定义，需要满足

$$\begin{pmatrix}O & I_{(m)}\\ -I_{(m)} & O\end{pmatrix}=H=\begin{pmatrix}I & X\\ O & I\end{pmatrix}^T H\begin{pmatrix}I & X\\ O & I\end{pmatrix}=\begin{pmatrix}O & I\\ -I & X^T-X\end{pmatrix},$$

于是，需要满足 $X^T = X$.

要使得 $\begin{pmatrix} I & O \\ X & I \end{pmatrix}$ **是辛方阵，依定义，需要满足**

$$\begin{pmatrix} O & I_{(m)} \\ -I_{(m)} & O \end{pmatrix} = H = \begin{pmatrix} I & O \\ X & I \end{pmatrix}^T H \begin{pmatrix} I & O \\ X & I \end{pmatrix} = \begin{pmatrix} X - X^T & I \\ -I & O \end{pmatrix},$$

于是，需要满足 $X^T = X$.

作业中的一些问题：

9.4. **第** 4 **题. 用正交方阵化二次型为标准形。**

一般过程就是将二次型写成矩阵形式，求其特征值，以及对应的特征向量，再进行 Gram-Schmidt **正交化程序把这些特征向量化成标准正交的。很多同学只进行到了求特征值这步。**

补充内容

1. **多线性映射，**Multilinear Mapping:

设 $V_1, \dots, V_s, W$ **为实数** $\mathbb{R}$ **上的线性空间。称一个映射**

$$f : \prod_{i=1}^{s} V_i \to W$$

为一个多线性映射，如果 f **对每一个分量都具有线性性，即对任意的** $1 \leqslant i \leqslant s$ **都有**

- $f(\alpha_1, \dots, \alpha_i + \alpha_i', \dots, \alpha_s) = f(\alpha_1, \dots, \alpha_i, \dots, \alpha_s) + f(\alpha_1, \dots, \alpha_i', \dots, \alpha_s)$
- $f(\alpha_1, \dots, t\alpha_i, \dots, \alpha_s) = tf(\alpha_1, \dots, \alpha_i, \dots, \alpha_s)$

类似于线性映射的情形，我们可以定义多线性映射的加法以及数乘，于是从 $\prod_{i=1}^{s} V_i$ **到** W **的多线性映射的全体成为了一个线性空间，被记作** $\mathcal{L}(V_1, \dots, V_s; W)$.

多线性映射的例子:

- **之前在学习方阵行列式的时候，已经提到过行列式** det **对于矩阵的列（以及行）有多线性性，于是我们可以把** det **视作是如下的多线性映射（函数）**

$$\det : \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n\text{ 个}} \to \mathbb{R}$$

- **内积也是多线性函数的一个例子**

$$\langle\ ,\ \rangle : V \times V \to \mathbb{R},$$

2. **张量积**:

设 $V_1, V_2, \dots, V_s$ **为线性空间。如果一个线性空间** T, **以及多线性映射** $\tau : \prod_{i=1}^{s} V_i \to T$ **满足如下的泛性质（**Universal Property**）**:

- **对任意线性空间** W **以及任意多线性映射** $f : \prod_{i=1}^{s} V_i \to W$, f **都能唯一地通过** τ **分解**:

$$\begin{tikzcd} V_1 \times \cdots \times V_s \arrow[r, "f"] \arrow[d, "\tau"'] & W \\ T \arrow[ur, dashed, "\exists! \tilde{f}"'] & \end{tikzcd}$$

那么 T **被称作** $V_1, V_2, \dots, V_s$ **的张量积（**Tensor Product**），记作** $V_1 \otimes V_2 \otimes \cdots \otimes V_s$. **直积空间** $\prod_{i=1}^{s} V_i$ **中的元素** $(\alpha_1, \dots, \alpha_s)$ **在** τ **下的像被记作** $\alpha_1 \otimes \cdots \otimes \alpha_s$.

张量有如下的性质:

- $\alpha_1 \otimes \cdots \otimes (\alpha_i + \alpha_i') \otimes \cdots \otimes \alpha_s = \alpha_1 \otimes \cdots \otimes \alpha_i \otimes \cdots \otimes \alpha_s + \alpha_1 \otimes \cdots \otimes \alpha_i' \otimes \cdots \otimes \alpha_s$

- $\alpha_1 \otimes \cdots \otimes (k\alpha_i) \otimes \cdots \otimes \alpha_s = k\alpha_1 \otimes \cdots \otimes \alpha_s$

为了记号方便，考虑 2 个线性空间的张量积 $V_1 \otimes V_2$.

在分别取定基 $\alpha_1, \ldots, \alpha_n$ 以及 $\beta_1, \ldots, \beta_m$ 之后，$V_1 \otimes V_2$ 的一组基就可以取作

$$\alpha_i \otimes \beta_j, \quad 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m$$

于是，$V_1 \otimes V_2$ 中的所有元素都能表示成有限和

$$\sum_{i,j} k_{ij} \alpha_i \otimes \beta_j.$$

于是，$\dim V_1 \otimes V_2 = \dim V_1 \times \dim V_2$.

注意！一般来说，$V_1 \otimes V_2$ 中元素并不能表示成 $\alpha \otimes \beta$, $\alpha \in V_1$, $\beta \in V_2$, 这种形式，例如 $\alpha_1 \otimes \beta_1 + \alpha_2 \otimes \beta_2$. 这点在物理中的重要应用就是用来表示量子纠缠态 (Quantum Entanglement).

在取定了 $V_i(i = 1, 2, \ldots, s)$ 的一组基

$$\varepsilon_{i1}, \ldots, \varepsilon_{in_i}$$

之后，$\bigotimes\limits_{i=1}^{s} V_i$ 中所有元素都能表示为一个和式

$$\sum_{i_1, i_2, \ldots, i_s} k_{i_1 i_2 \ldots i_s} \varepsilon_{1i_1} \otimes \cdots \otimes \varepsilon_{si_s}$$

其中 $1 \leqslant i_1 \leqslant \dim V_1 = n_1, \ldots, 1 \leqslant i_s \leqslant \dim V_s = n_s$.

张量除了加法与数乘，还可以定义他们的积。为了书写方便，我们任取两个张量空间 $V_1 \otimes \cdots \otimes V_s$ 与 $W_1 \otimes \cdots \otimes W_t$ 中的两个张量

$$v = \alpha_1 \otimes \cdots \otimes \alpha_s, \quad w = \beta_1 \otimes \cdots \otimes \beta_t$$

那么可以定义 v 与 w 的积为

$$v \otimes w := \alpha_1 \otimes \cdots \otimes \alpha_s \otimes \beta_1 \otimes \cdots \otimes \beta_t$$

为张量空间 $V_1 \otimes \cdots \otimes V_s \otimes W_1 \otimes \cdots \otimes W_t$ 中的元素。

大家可以自行写一写以上两个张量空间中的一般元素

$$\sum k_{i_1, i_2, \ldots, i_s} \alpha_{1i_1} \otimes \cdots \otimes \alpha_{si_s}, \quad \sum h_{j_1, j_2, \ldots, j_t} \beta_{1j_1} \otimes \cdots \otimes \beta_{tj_t}$$

的积的表达式。

作为多维数组的张量:

考虑张量空间 $\bigotimes\limits_{i=1}^{s} V_i$ 中一般的一个元素

$$\sum_{i_1, i_2, \ldots, i_s} k_{i_1 i_2 \ldots i_s} \varepsilon_{1i_1} \otimes \cdots \otimes \varepsilon_{si_s}$$

其中 $\varepsilon_{i1}, \ldots, \varepsilon_{in_i}, i = 1, 2, \ldots, s$, 为线性空间 V_i 的一组基，且令 $n_i = \dim V_i$. 那么该元素能够被视作一个 s 维数组

`double` $k[n_1][n_2]...[n_s]$;

使得数组中 $k[i_1][i_2]...[i_s]$ 的值就是 $k_{i_1 i_2 \ldots i_s}$.