DS4400 HW1

Xin Guan

1. Let $a \in \mathbb{R}^n$ be an n-dimensional vector and let $\boldsymbol{U} \in \mathbb{R}^{n\times n}$ be an orthonormal matrix, i.e., $\boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^T = \mathbf{I}_n$. show the following:

   (a) $\text{trace}(\boldsymbol{a}\boldsymbol{a}^T) = \|\boldsymbol{a}\|_2^2$

   Solution:

   Proof. Let $\boldsymbol{a} = < a_1, a_2, a_3 \dots a_n >$

   Then, $\boldsymbol{a}\boldsymbol{a}^T = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \cdot \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} = \begin{bmatrix} a_1 \cdot a_1 & a_1 \cdot a_2 & \dots & a_1 \cdot a_n \\ a_2 \cdot a_1 & a_2 \cdot a_2 & \dots & a_2 \cdot a_n \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_n \cdot a_1 & a_n \cdot a_2 & \dots & a_n \cdot a_n \end{bmatrix}$

   Therefore, $\text{trace}(\boldsymbol{a}\boldsymbol{a}^T) = a_1 \cdot a_1 + a_2 \cdot a_2 + \cdots + a_n \cdot a_n = \sum_{i=1}^n a_i^2 = \|\boldsymbol{a}\|_2^2$

   Thus, $\text{trace}(\boldsymbol{a}\boldsymbol{a}^T) = \|\boldsymbol{a}\|_2^2$ □

   (b) $\|\boldsymbol{U}\boldsymbol{a}\|_2^2 = \|\boldsymbol{a}\|_2^2$

   Solution:

   Proof. We can write $\boldsymbol{U}$ as follow:

   $\boldsymbol{U} = \begin{bmatrix} —v_1^T— \\ —v_2^T— \\ \dots \\ —v_n^T— \end{bmatrix}$, where $v_i \in \mathbb{R}^n, i \in \mathbf{Z}, 1 \leq i \leq n$.

   Since $\boldsymbol{U}$ is orthonormal, we have:

   $\forall i \neq k, v_i * v_k = \overrightarrow{\mathbf{0}}$

   $\forall i \in \mathbf{Z}, 1 \leq i \leq n, v_i^T v_i = 1$

   Then, we can write $\boldsymbol{U}\boldsymbol{a} = \begin{bmatrix} v_1^T\boldsymbol{a} \\ v_2^T\boldsymbol{a} \\ \dots \\ v_n^T\boldsymbol{a} \end{bmatrix}$

   Therefore, $\|\boldsymbol{U}\boldsymbol{a}\| = (v_1^T\boldsymbol{a})^2 + (v_2^T\boldsymbol{a})^2 + \cdots + (v_n^T\boldsymbol{a})^2$

   $= \sum_{i=1}^n (v_i^T\boldsymbol{a}) = \sum_{i=1}^n (v_i^T\boldsymbol{a})^T(v_i^T\boldsymbol{a}) = \sum_{i=1}^n (\boldsymbol{a}^T v_i)(v_i^T\boldsymbol{a})$

   $= \sum_{i=1}^n (\boldsymbol{a}^T(v_i v_i^T)\boldsymbol{a})$

   Since $\forall i \in \mathbf{Z}, 1 \leq i \leq n, v_i^T v_i = 1$, we have:

   $\sum_{i=1}^n (\boldsymbol{a}^T(v_i v_i^T)\boldsymbol{a}) = \sum_{i=1}^n (\boldsymbol{a}^T\boldsymbol{a}) = \|\boldsymbol{a}\|_2^2$

   Thus, $\|\boldsymbol{U}\boldsymbol{a}\|_2^2 = \|\boldsymbol{a}\|_2^2$. □

2. Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{B} \in \mathbb{R}^{n\times n}$ be arbitrary but invertible matrices and let α be a scalar. Show the following:

   (a) $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$

   Solution:

   Proof. By the definition of inverse, we have:

   $$(\boldsymbol{AB})(\boldsymbol{AB})^{-1} = \boldsymbol{I}_n$$

   Multiply both sides by $\boldsymbol{A}^{-1}$:

   $$\boldsymbol{A}^{-1}\boldsymbol{AB}(\boldsymbol{AB})^{-1} = \boldsymbol{A}^{-1}\boldsymbol{I}_n$$

   Then we have:

   $$\boldsymbol{B}(\boldsymbol{AB})^{-1} = \boldsymbol{A}^{-1}$$

   Multiply both sides by $\boldsymbol{B}^{-1}$:

   $$\boldsymbol{B}^{-1}\boldsymbol{B}(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$$

   Then we have:

   $$(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$$

   Therefore, $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$ □

   (b) $(\boldsymbol{A}^T)^{-1} = (\boldsymbol{A}^{-1})^T$

   Solution:

   Proof. By the definition of inverse, we have $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}_n$. Then we transpose both sides:

   $$(\boldsymbol{A}^{-1}\boldsymbol{A})^T = (\boldsymbol{I}_n)^T$$

   Then we have:

   $$\boldsymbol{A}^T(\boldsymbol{A}^{-1})^T = \boldsymbol{I}_n$$

   Multiply both sides by $(\boldsymbol{A}^T)^{-1}$:

   $$(\boldsymbol{A}^T)^{-1}\boldsymbol{A}^T(\boldsymbol{A}^{-1})^T = (\boldsymbol{A}^T)^{-1}\boldsymbol{I}_n$$

   Then we have:

   $$(\boldsymbol{A}^{-1})^T = (\boldsymbol{A}^T)^{-1}$$

   □

   (c) $\mathrm{trace}(\alpha\boldsymbol{A}) = \alpha\ \mathrm{trace}(\boldsymbol{A})$

   Solution:

   Proof. Let $\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \ddots & & \vdots \\ a_{11} & a_{12} & \dots & a_{1n} \end{bmatrix}$

   Then $\alpha\boldsymbol{A} = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \dots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \dots & \alpha a_{2n} \\ \vdots & \ddots & & \vdots \\ \alpha a_{11} & \alpha a_{12} & \dots & \alpha a_{nn} \end{bmatrix}$

   Then, $\mathrm{trace}(\alpha\boldsymbol{A}) = \alpha a_{11} + \alpha a_{22} + \dots \alpha a_{nn} = \sum_{i=1}^{n} \alpha a_{ii} = \alpha \sum_{i=1}^{n} a_{ii}$

   On the other hand, $\alpha\ \mathrm{trace}(\boldsymbol{A}) = \alpha \cdot (a_{11} + a_{22} + \dots a_{nn}) = \alpha \sum_{i=1}^{n} a_{ii}$

   Therefore, $\mathrm{trace}(\alpha\boldsymbol{A}) = \alpha\ \mathrm{trace}(\boldsymbol{A})$ □