EE2100: Matrix Analysis

Review Notes - 26

Topics covered :

1. Trace of a matrix.
2. Axiomatic definition of determinant.

---

1. The Trace of a matrix, which is defined only for square matrices, is the sum of the diagonal entries of the matrix. The trace of a square matrix (say $\mathbf{A}$) is denoted by $\mathbf{Tr}\left(\mathbf{A}\right)$. It can be thought of as a function that operates on a matrix and gives a number and is defined as

$$\mathbf{Tr}\left(\mathbf{A}\right) : \mathcal{R}^{n\times n} \rightarrow \mathcal{R} = \sum_{i=1}^{n} A_{ii} \tag{1}$$

The function Trace satisfies the following properties.

(a) $\mathbf{Tr}(\alpha\mathbf{A}) = \alpha\mathbf{Tr}(\mathbf{A})$.

(b) $\mathbf{Tr}(\mathbf{A}+\mathbf{B}) = \mathbf{Tr}(\mathbf{A}) + \mathbf{Tr}(\mathbf{B})$.

(c) The Trace is a linear function.

2. Let $\mathbf{C} = \mathbf{A}^T\mathbf{B}$ where $\mathbf{A} \in \mathcal{R}^{n\times n}$ and $\mathbf{B} = \mathcal{R}^{n\times n}$. The trace of $\mathbf{C}$ is the sum of dot product between the corresponding column vectors of $\mathbf{A}$ and $\mathbf{B}$.

3. The trace of matrix $\mathbf{C} = \mathbf{A}^T\mathbf{A}$ gives the square of the Frobenius norm of the matrix $\mathbf{A}$.

4. Let $\mathbf{A}$ and $\mathbf{B}$ be two matrices such that $\mathbf{AB}$ and $\mathbf{BA}$ are defined and their product results in a $n \times n$ matrix. Then $\mathbf{Tr}(\mathbf{AB}) = \mathbf{Tr}(\mathbf{BA})$ (Proof part of the assignment).

5. Axiomatic definition of trace: The trace of a matrix (typically denoted by $\mathbf{Tr}(\mathbf{A})$), defined for square matrices, is a function (say f, for analysis) that satisfies the following conditions

(a) Linearity: Let $\mathbf{A}$,$\mathbf{B} \in \mathcal{R}^{n\times n}$. The function is linear if $f\left(\alpha\mathbf{A}+\beta\mathbf{B}\right) = \alpha f\left(\mathbf{A}\right) + \beta f\left(\mathbf{B}\right)$ for $\alpha, \beta \in \mathcal{R}$.

(b) Let $\mathbf{A}$,$\mathbf{B} \in \mathcal{R}^{n\times n}$. Then $f\left(\mathbf{AB}\right) = f\left(\mathbf{BA}\right)$.

(c) $f\left(\mathbf{I}\right) = n$ where $\mathbf{I} \in \mathcal{R}^{n\times n}$.

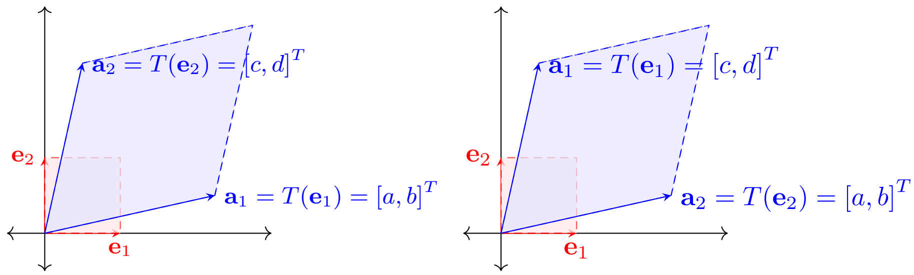
6. Geometric interpretation of Determinant: Let $\mathbf{A} \in \mathcal{R}^{2\times 2}$ whose entries are

$$\mathbf{A} = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \tag{2}$$

It is known from the earlier courses (at pre-university level) that the determinant of the 2×2 matrix is given by $\mathbf{det}(\mathbf{A}) = ad - bc$ (We'll derive this same equation in Lecture 31). The aim here is to understand the significance of the determinant.

If we consider the matrix to represent a linear transformation (say T), we can infer the first column vector (i.e., $\mathbf{a}_1$) as the output of transformation applied to $\mathbf{e}_1$ (see Fig. 1(a)) and the second column vector (i.e., $\mathbf{a}_2$) as the output of the transformation applied to $\mathbf{e}_2$ (see Fig. 1(a)).

The parallelogram (rectangle to be precise; shown in Fig. 1(a)) formed by the vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ has an area of 1. On the other hand, the parallelogram formed by $\mathbf{a}_1$ and $\mathbf{a}_2$ has an area of $ad - bc$ (can be derived in several ways).



(a) Result of transformation T (b) Result of transformation related to $\mathbf{B}$

Fig. 1: Geometric interpretation of the determinant

So it can be inferred that the determinant represents the ratio in which the area of the parallelogram formed by $\mathbf{e}_1$ and $\mathbf{e}_2$ would scale when a linear transformation T is applied to it. Let B denote a matrix whose entries are

$$\mathbf{B} = \begin{bmatrix} c & a \\ d & b \end{bmatrix} \tag{3}$$

The determinant in this case is $bc - ad$ and is equal to the negative of the area of the transformed parallelogram (see Fig. 1(b)).

In general, the magnitude of the determinant of a 2×2 matrix gives the factor by which the areas of the parallelogram get transformed when subjected to the associated linear transformation.

For the higher dimensional case, the magnitude of the determinant would indicate the ratios in which the volume of the parallelepiped would scale when subjected to the associated linear transformation.

7. Axiomatic definition of determinant: The determinant of a matrix (typically denoted by $\mathbf{det}(\mathbf{A})$), defined for square matrices, is a function that satisfies the following conditions

   (a) Multilinearity: Let $\mathbf{A} \in \mathcal{R}^{n\times n}$ and let $f(\mathbf{A})$ denote the determinant of $\mathbf{A}$. To understand multilinearity, we analyze determinant as a function that operates on column vectors of $\mathbf{A}$ i.e. as $f(\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n)$. We say that the function is multilinear if $f(\mathbf{a}_1, \mathbf{a}_2, \cdots, \alpha\mathbf{u}_i + \beta\mathbf{v}_i, \cdots, \mathbf{a}_n) = \alpha f(\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{u}_i, \cdots, \mathbf{a}_n) + \beta f(\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{v}_i, \cdots, \mathbf{a}_n)$.

   (b) Alternating property: Let $f(\mathbf{A}) = f(\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n)$ denote the determinant of $\mathbf{A}$, then $f(\mathbf{a}_2, \mathbf{a}_1, \cdots, \mathbf{a}_n) = -f(\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n)$ i.e., the determinant should change in sign if two columns are interchanged.

   (c) Unity for Identity matrix: The determinant of an identity matrix is unity.