Lecture Note on Linear Algebra

12. More about Determinants

Wei-Shi Zheng,
wszheng@ieee.org，2011

October 24, 2011

1 What Do You Learn from This Note

In the last lecture, we have introduced the concept of determinant, but we should ask ourselves the following questions (I think many of you are already on this right track):

1. Why do we need to formulate determinant?
2. How can it be used for solving a matrix equation?
3. Any geometric interpretation about determinant?

We are now going to address all these questions.

注：对于理解上面问题的第三，现在看来有难度，所以本lecture note的重点是掌握前两点，了解第3点。

Basic concept：adjugate of A(矩阵A的共轭)

2 Adjugate of A(矩阵A的共轭)

We first clarify a notation. Let $A = (\vec{a}_1 \ \cdots \ \vec{a}_i \ \cdots \ \vec{a}_n)$ be a $n \times n$ square matrix and $\vec{b} \in \mathbb{R}^n$. Then write $A_i(\vec{b})$ for the matrix $(\vec{a}_1 \ \cdots \ \vec{b} \ \cdots \ \vec{a}_n)$, that is, $A_i(\vec{b})$ is resulted from A by replacing the i–th column of A by $\vec{b}$.

2.1 A formula for the inverse of a square matrix

Let A be the $n \times n$ square matrix. Define adj$A$, the adjugate of A(矩阵A的共轭), to be the square matrix such that $[\text{adj}A]_{\underline{ij}} = C_{ji}$ (recall that C_{ji} is the (j,i)–th cofactor of A), i.e. $C_{ji} = (-1)^{j+i} \det \overline{A}_{ji}$. That is:

$$\text{adj}A = \begin{pmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{pmatrix} \text{大家这里小心看看}C\text{的下标分布.}$$

Let us compute the product (adj$A$)A. We have

$$[(\text{adj}A)A]_{ij} = \sum_{k=1}^{n} [\text{adj}A]_{ik} a_{kj} = \sum_{k=1}^{n} a_{kj} C_{ki} = \det A_i(\vec{a}_j) = \begin{cases} \det A & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

That is

$$(\text{adj}A)A = \begin{pmatrix} \det A & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \det A \end{pmatrix} = (\det A) I_n.$$

Similarly, $A\text{adj}A = (\det A)I_n$. So for the case that A is invertible, we have

$$\left(\frac{1}{\det A}(\text{adj}A)\right) A = A \left(\frac{1}{\det A}(\text{adj}A)\right) = I_n.$$

Thus we obtain that $A^{-1} = \frac{1}{\det A}(\text{adj}A)$, which is a formula for computing A^{-1}.

Remark: This formula is used only in theoretical study in mathematics. It is impractical to use it in practical computation of A^{-1}, in which we normally use the method introduced in Linear Algebra 8.

Example: Textbook P.203.

2.2 Cramer's Rule

We use the formula derived in the last section to give another formula for the solution of $A\vec{x} = \vec{b}$ where A is a $n \times n$ matrix and is invertible, and $\vec{b} \in \mathbb{R}^n$. Since A is invertible, we know that the equation $A\vec{x} = \vec{b}$ has a unique solution, namely $vecx = A^{-1}\vec{b}$. So

$$\vec{x} = A^{-1}\vec{b} = \frac{1}{\det A}(\text{adj}A)\vec{b}.$$

Consider $(\text{adj}A)\vec{b}$, we have

$$[(\text{adj}A)\vec{b}]_i = \sum_{j=1}^{n}[\text{adj}A]_{ij}b_j = \sum_{j=1}^{n} b_j C_{ji} = \det A_i(\vec{b}),$$

which leads to

$$x_i = \frac{[(\text{adj}A)\vec{b}]_i}{\det A} = \frac{\det A_i(\vec{b})}{\det A}.$$

Thus we obtain the so called Cramer's Rule:

$$\vec{x} = A^{-1}\vec{b} = \frac{1}{\det A}\begin{pmatrix} \det A_1(\vec{b}) \\ \vdots \\ \det A_n(\vec{b}) \end{pmatrix},$$

which is only useful in theoretical study but not practical computation.

注：上面定理也就是说把b替换矩阵A的每一列，然后分别计算这些矩阵的行列式，并把这些行列式的值按顺序构建出向量，最后对该向量除$\det A$即是矩阵方程的解。

2.3 Determinants as Area(面积) and Volume(体积)

注：本节理解难度较大，以了解为主。

We shall give a geometric interpretation on determinants of matrices.

THEOREM 1.
Let $A = (\vec{a}_1\ \vec{a}_2)$ be a 2×2 square matrix and $S(A)$ denote the area of the parallelogram determined by vertices $\vec{0}$, $\vec{a}_1$, $\vec{a}_2$. Then

$$S(A) = |\det A|.$$

Proof. The proof is again done by factorizing A into elementary matrices.

1. A is not invertible: Then $\vec{a}_1$, $\vec{a}_2$ are linearly dependent. The parallelogram is degenerated to a line segment or a point. In this case, $S(A) = |\det A| = 0$.

2. A is invertible. Firstly, it is easy to verify that for any 2×2 square matrix A and any elementary matrix E of size 2, we have

   $$S(EA) = |\det E| S(A).$$

   (见板书解释和lecture note 11 定理7)
   Now A is invertible means $A = E_l \cdots E_1$ where $E_1, \ldots, E_l$ are elementary matrices. So

   $$\begin{aligned} S(A) &= S(E_l \cdots E_1 \cdot I_2) \\ &= S(E_{l-1} \cdots E_1 \cdot I_2) |\det E_l| \\ &\cdots \\ &= |\det E_l| \cdots |\det E_1| S(I_2) \\ &= |\det E_l| \cdots |\det E_1| \quad [\text{Obviously, } S(I_2) = 1] \\ &= |\det E_l \cdots \det E_1| \\ &= |\det E_l \cdots E_1| \\ &= |\det A|. \end{aligned}$$

□

Example: Find the area of the parallelogram with vertices $\vec{a}_0$, $\vec{a}_1$, $\vec{a}_2$.
Solution $S = S(\vec{a}_1 - \vec{a}_0\ \vec{a}_2 - \vec{a}_0) = |\det(\vec{a}_1 - \vec{a}_0\ \vec{a}_2 - \vec{a}_0)|$.

For parallelepiped in $\mathbb{R}^3$, we have

THEOREM 2.
Let $A = (\vec{a}_1\ \vec{a}_2\ \vec{a}_3)$ be a 3×3 matrix and $V(A)$ denote the volume of the parallelepiped determined by vertices $\vec{0}$, $\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$. Then

$$V(A) = |\det A|.$$

The proof of THEOREM 2 is almost identical to that of THEOREM 1. We shall omit it here.

Remark: The result of THEOREM 1 can be generalized to $\mathbb{R}^n$ for any $n \in \mathbb{Z}^+$. However, the concept of 'volumn' (or Lebesgue measure formally) in higher dimensions need to be clarified in general. This is a big problem. Detailed discussion will be found in Measure Theory.

2.4 Transformation and determinant

The determinant of the standard matrix for a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ also gives a geometric property of T. We explain this fact as follows.

Let K denote a subset of $\mathbb{R}^2$ for which the area of K is well–defined. (e.g. region bounded by triangle, parallelogram, circle, etc.). Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be linear with standard matrix A. Then the area of the image $T(K)$ of K is also well–defined and satisfies

$$S(T(K)) = |\det A| S(K).$$

The general proof of this result is out of the scope of this course. However, the special case where K is the region bounded by the parallelogram determined by vertices $\vec{a}_0$, $\vec{a}_1$ and $\vec{a}_2$ can be shown using the result of Theorem 1 in the last section. Let us detail it in the following.

We obverse that the image $T(K)$ of K under T is also a region bounded by the parallelogram determined by vertices $T(\vec{a}_0)$, $T(\vec{a}_1)$ and $T(\vec{a}_2)$. So

$$\begin{aligned}
S(T(K)) &= S(T(\vec{a}_1) - T(\vec{a}_0), T(\vec{a}_2) - T(a_0)) \\
&= S(A\vec{a}_1 - A\vec{a}_0, A\vec{a}_2 - A\vec{a}_0) \\
&= S(A(\vec{a}_1 - \vec{a}_0, \vec{a}_2 - \vec{a}_0)) \\
&= |\det A(\vec{a}_1 - \vec{a}_0, \vec{a}_2 - \vec{a}_0)| \\
&= |\det A| |\det(\vec{a}_1 - \vec{a}_0, \vec{a}_2 - \vec{a}_0)| \\
&= |\det A| S(K).
\end{aligned}$$

Again this result can be generalised to any higher dimension. For instance, for any subset K in $\mathbb{R}^3$ for which the volume of K is well–defined. Then

$$V(T(K)) = |\det A| V(K),$$

where T is any linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^3$ and A the standard matrix for T. Again, the proof is out of scope.

THE TOWER OF BABEL (通天塔), *by Bruegel the elder*