MATH 323

Linear Algebra

Lecture 6: Matrix algebra (continued). Determinants.

General results on inverse matrices

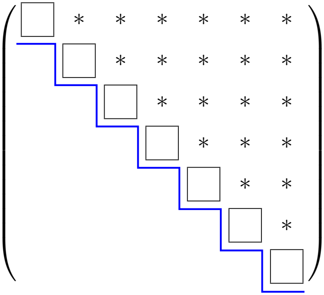
Theorem 1 Given an $n \times n$ matrix A, the following conditions are equivalent:

(i) A is invertible;
(ii) $\mathbf{x} = \mathbf{0}$ is the only solution of the matrix equation $A\mathbf{x} = \mathbf{0}$;
(iii) the matrix equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for any n-dimensional column vector $\mathbf{b}$;
(iv) the row echelon form of A has no zero rows;
(v) the reduced row echelon form of A is the identity matrix.

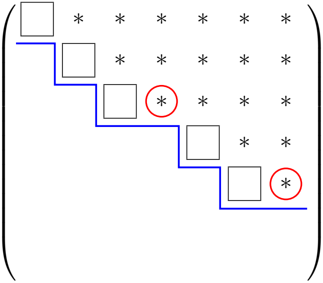
Theorem 2 Suppose that a sequence of elementary row operations converts a matrix A into the identity matrix.

Then the same sequence of operations converts the identity matrix into the inverse matrix A^{-1}.

Row echelon form of a square matrix:



invertible case



noninvertible case

Why does it work?

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ 2b_1 & 2b_2 & 2b_3 \\ c_1 & c_2 & c_3 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1+3a_1 & b_2+3a_2 & b_3+3a_3 \\ c_1 & c_2 & c_3 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ c_1 & c_2 & c_3 \\ b_1 & b_2 & b_3 \end{pmatrix}.$$

Theorem Any elementary row operation can be simulated as left multiplication by a certain matrix (called an *elementary matrix*).

Elementary matrices

$$E = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & O & \\ & & 1 & & & & \\ & & & r & & & \\ & & & & 1 & & \\ & O & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \quad \text{row } \#i$$

To obtain the matrix EA from A, multiply the ith row by r. To obtain the matrix AE from A, multiply the ith column by r.

Elementary matrices

$$E = \begin{pmatrix} 1 & & & & & & \\ \vdots & \ddots & & & & O & \\ 0 & \cdots & 1 & & & & \\ \vdots & & \vdots & \ddots & & & \\ 0 & \cdots & r & \cdots & 1 & & \\ \vdots & & \vdots & & \vdots & \ddots & \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} \begin{matrix} \\ \\ \text{row } \#i \\ \\ \text{row } \#j \\ \\ \\ \end{matrix}$$

To obtain the matrix EA from A, add r times the ith row to the jth row. To obtain the matrix AE from A, add r times the jth column to the ith column.

Elementary matrices

$$E = \begin{pmatrix} 1 & & & & & O & \\ & \ddots & & & & & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & \ddots & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & O & & & & & 1 \end{pmatrix} \quad \begin{matrix} \\ \\ \text{row } \#i \\ \\ \text{row } \#j \\ \\ \\ \end{matrix}$$

To obtain the matrix EA from A, interchange the ith row with the jth row. To obtain AE from A, interchange the ith column with the jth column.

Elementary matrices

Theorem 1 Any elementary row operation σ on matrices with n rows can be simulated as left multiplication by a certain $n \times n$ matrix E_σ (called an *elementary matrix*).

Theorem 2 Elementary matrices are invertible.

Proof: Suppose E_σ is an $n \times n$ elementary matrix corresponding to an operation σ. We know that σ can be undone by another elementary row operation τ. It is easy to check that σ undoes τ as well. Then for any matrix A with n rows we have $E_\tau E_\sigma A = A$ (since τ undoes σ) and $E_\sigma E_\tau A = A$ (since σ undoes τ). In particular, $E_\tau E_\sigma I = E_\sigma E_\tau I = I$, which implies that $E_\tau = E_\sigma^{-1}$.

Theorem 3 A square matrix is invertible if and only if it can be expanded into a product of elementary matrices.

Theorem Suppose that a sequence of elementary row operations converts a matrix A into the identity matrix.

Then the same sequence of operations converts the identity matrix into the inverse matrix A^{-1}.

Proof: Let $E_1, E_2, \dots, E_k$ be elementary matrices that correspond to elementary row operations converting A into I. Then $E_k E_{k-1} \dots E_2 E_1 A = I$.

Applying the same sequence of operations to the identity matrix I, we obtain the matrix

$$B = E_k E_{k-1} \dots E_2 E_1 I = E_k E_{k-1} \dots E_2 E_1.$$

Therefore $BA = I$. Besides, B is invertible since elementary matrices are invertible. Then $B^{-1}(BA) = B^{-1}I$. It follows that $A = B^{-1}$, hence $B = A^{-1}$.

Theorem A square matrix A is invertible if and only if $\mathbf{x} = \mathbf{0}$ is the only solution of the matrix equation $A\mathbf{x} = \mathbf{0}$.

Corollary 1 For any $n \times n$ matrices A and B,

$$BA = I \iff AB = I.$$

Proof: It is enough to prove that $BA = I \implies AB = I$. Assume $BA = I$. Then $A\mathbf{x} = \mathbf{0} \implies B(A\mathbf{x}) = B\mathbf{0}$ $\implies (BA)\mathbf{x} = \mathbf{0} \implies \mathbf{x} = \mathbf{0}$. By the theorem, A is invertible. Then $BA = I \implies A(BA)A^{-1} = AIA^{-1} \implies AB = I$.

Corollary 2 Suppose A and B are $n \times n$ matrices. If the product AB is invertible, then both A and B are invertible.

Proof: Let $C = B(AB)^{-1}$ and $D = (AB)^{-1}A$. Then $AC = A\big(B(AB)^{-1}\big) = (AB)(AB)^{-1} = I$ and $DB = \big((AB)^{-1}A\big)B = (AB)^{-1}(AB) = I$. By Corollary 1, $C = A^{-1}$ and $D = B^{-1}$.

Transpose of a matrix

Definition. Given a matrix A, the **transpose** of A, denoted A^T, is the matrix whose rows are columns of A (and whose columns are rows of A). That is, if $A = (a_{ij})$ then $A^T = (b_{ij})$, where $b_{ij} = a_{ji}$.

Examples. $$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix},$$

$$\begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}^T = (7, 8, 9), \qquad \begin{pmatrix} 4 & 7 \\ 7 & 0 \end{pmatrix}^T = \begin{pmatrix} 4 & 7 \\ 7 & 0 \end{pmatrix}.$$

Properties of transposes:

- $(A^T)^T = A$
- $(A + B)^T = A^T + B^T$
- $(rA)^T = rA^T$
- $(AB)^T = B^T A^T$
- $(A_1 A_2 \dots A_k)^T = A_k^T \dots A_2^T A_1^T$
- $(A^{-1})^T = (A^T)^{-1}$

Definition. A square matrix A is said to be **symmetric** if $A^T = A$.

For example, any diagonal matrix is symmetric.

Proposition For any square matrix A the matrices $B = AA^T$ and $C = A + A^T$ are symmetric.

Proof:

$$B^T = (AA^T)^T = (A^T)^T A^T = AA^T = B,$$

$$C^T = (A + A^T)^T = A^T + (A^T)^T = A^T + A = C.$$

Determinants

Determinant is a scalar assigned to each square matrix.

Notation. The determinant of a matrix $A = (a_{ij})_{1\le i,j\le n}$ is denoted $\det A$ or

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Principal property: $\det A \neq 0$ if and only if a system of linear equations with the coefficient matrix A has a unique solution. Equivalently, $\det A \neq 0$ if and only if the matrix A is invertible.

Definition in low dimensions

Definition. $\det(a) = a, \quad \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc,$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}.$$

$$+: \quad \begin{pmatrix} \boxed{*} & * & * \\ * & \boxed{*} & * \\ * & * & \boxed{*} \end{pmatrix}, \begin{pmatrix} * & \boxed{*} & * \\ * & * & \boxed{*} \\ \boxed{*} & * & * \end{pmatrix}, \begin{pmatrix} * & * & \boxed{*} \\ \boxed{*} & * & * \\ * & \boxed{*} & * \end{pmatrix}.$$

$$-: \quad \begin{pmatrix} * & * & \boxed{*} \\ * & \boxed{*} & * \\ \boxed{*} & * & * \end{pmatrix}, \begin{pmatrix} * & \boxed{*} & * \\ \boxed{*} & * & * \\ * & * & \boxed{*} \end{pmatrix}, \begin{pmatrix} \boxed{*} & * & * \\ * & * & \boxed{*} \\ * & \boxed{*} & * \end{pmatrix}.$$

Examples: 2×2 matrices

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1, \qquad \begin{vmatrix} 3 & 0 \\ 0 & -4 \end{vmatrix} = -12,$$

$$\begin{vmatrix} -2 & 5 \\ 0 & 3 \end{vmatrix} = -6, \qquad \begin{vmatrix} 7 & 0 \\ 5 & 2 \end{vmatrix} = 14,$$

$$\begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} = 1, \qquad \begin{vmatrix} 0 & 0 \\ 4 & 1 \end{vmatrix} = 0,$$

$$\begin{vmatrix} -1 & 3 \\ -1 & 3 \end{vmatrix} = 0, \qquad \begin{vmatrix} 2 & 1 \\ 8 & 4 \end{vmatrix} = 0.$$

Examples: 3×3 matrices

$$\begin{vmatrix} 3 & -2 & 0 \\ 1 & 0 & 1 \\ -2 & 3 & 0 \end{vmatrix} = 3\cdot 0\cdot 0 + (-2)\cdot 1\cdot(-2) + 0\cdot 1\cdot 3$$

$$- 0\cdot 0\cdot(-2) - (-2)\cdot 1\cdot 0 - 3\cdot 1\cdot 3 = 4 - 9 = -5,$$

$$\begin{vmatrix} 1 & 4 & 6 \\ 0 & 2 & 5 \\ 0 & 0 & 3 \end{vmatrix} = 1\cdot 2\cdot 3 + 4\cdot 5\cdot 0 + 6\cdot 0\cdot 0$$

$$- 6\cdot 2\cdot 0 - 4\cdot 0\cdot 3 - 1\cdot 5\cdot 0 = 1\cdot 2\cdot 3 = 6.$$

General definition

The general definition of the determinant is quite complicated as there is no simple explicit formula.

There are several approaches to defining determinants.

Approach 1 (original): an explicit (but very complicated) formula.

Approach 2 (axiomatic): we formulate properties that the determinant should have.

Approach 3 (inductive): the determinant of an $n \times n$ matrix is defined in terms of determinants of certain $(n-1)\times(n-1)$ matrices.

Classical definition

Definition. If $A = (a_{ij})$ is an $n \times n$ matrix then

$$\det A = \sum_{\pi \in S_n} \operatorname{sgn}(\pi)\, a_{1,\pi(1)}\, a_{2,\pi(2)} \dots a_{n,\pi(n)},$$

where π runs over S_n, the set of all permutations of $\{1, 2, \dots, n\}$, and $\operatorname{sgn}(\pi)$ denotes the sign of the permutation π.

Remarks. • A **permutation** of the set $\{1, 2, \dots, n\}$ is an invertible mapping of this set onto itself. There are $n!$ such mappings.

• The **sign** $\operatorname{sgn}(\pi)$ can be 1 or -1. Its definition is rather complicated.