Theorem 2.8. *Given a square matrix A, let use reduce it to echelon form B by elementary row operations of types (1) and (2). If B has a zero row, then $\det A = 0$. Otherwise, let k be the number of row exchanges involved in the reduction process. Then $\det A$ equals $(-1)^k$ times the product of the diagonal entries of B.*

Definition. Let A be a matrix of size n by m; let B and C be matrices of size m by n. We say that B is a **left inverse** for A if $B \cdot A = I_m$, and we say that C is a **right inverse** for A if $A \cdot C = I_n$.

Theorem 2.2. *If A has both a left inverse B and a right inverse C, then they are unique and equal.*

Theorem 2.7. *Let A be a square matrix. If the rows of A are independent, then $\det A \neq 0$; if the rows are dependent, then $\det A = 0$. Thus an n by n matrix A has rank n if and only if $\det A \neq 0$.*

Definition. Let A be an n by n matrix. The matrix of size $n-1$ by $n-1$ that is obtained from A by deleting the i^{th} row and the j^{th} column of A is called the **(i,j)-minor** of A. It is denoted A_{ij}. The number

$$(-1)^{i+j} \det A_{ij}$$

Theorem 2.5. *If A is a square matrix and if B is a left inverse for A, then B is also a right inverse for A.*

Definition. A function that assigns, to each n by n matrix A, a real number denoted $\det A$, is called a **determinant function** if it satisfies the following axioms:

(1) If B is the matrix obtained by exchanging any two rows of A, then $\det B = -\det A$.

(2) Given i, the function $\det A$ is linear as a function of the i^{th} row alone.

(3) $\det I_n = 1$.

Theorem 2.4. *Let A be a matrix of size n by m. Suppose*

$$n = m = \text{rank } A.$$

Then A is invertible; and furthermore, A equals a product of elementary matrices.

Theorem 2.14. *Let A be an n by n matrix of rank n; let $B = A^{-1}$. Then*

$$b_{ij} = \frac{(-1)^{j+i} \det A_{ji}}{\det A}.$$

Theorem 2.1. *Let A be an n by m matrix. Any elementary row operation on A may be carried out by premultiplying A by the corresponding elementary matrix.*

Theorem 2.6. *Let A be an n by n matrix.*

(a) *If E is the elementary matrix corresponding to the operation that exchanges rows i_1 and i_2, then $\det(E \cdot A) = -\det A$.*

(b) *If E' is the elementary matrix corresponding to the operation that replaces row i_1 of A by itself plus c times row i_2, then $\det(E' \cdot A) = \det A$.*

(c) *If E'' is the elementary matrix corresponding to the operation that multiplies row i of A by the non-zero scalar λ, then $\det(E'' \cdot A) = \lambda(\det A)$.*

(d) *If A is the identity matrix I_n, then $\det A = 1$.*

Theorem 2.3. *Let A be a matrix of size n by m. If A is invertible, then*

$$n = m = \text{rank } A.$$

Definition. If A has both a right inverse and a left inverse, then A is said to be **invertible**. The unique matrix that is both a right inverse and a left inverse for A is called the **inverse** of A, and is denoted A^{-1}.

Theorem 2.15. *Let A be an n by n matrix. Let i be fixed. Then*

$$\det A = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \cdot \det A_{ik}.$$

$$E'' = \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \lambda & & & \\ & & & \ddots & & \\ & & & & 1 & \\ & & & & & 1 \end{bmatrix} \nwarrow \text{row } i.$$

$$E = \begin{bmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & 1 & \\ & & & & & & 1 \end{bmatrix} \begin{matrix} \nwarrow \text{row } i_1 \\ \swarrow \text{row } i_2 \end{matrix}.$$

$$E' = \begin{bmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & 1 & \cdots & c & & \\ & & \vdots & & \vdots & & \\ & & 0 & \cdots & 1 & & \\ & & & & & 1 & \\ & & & & & & 1 \end{bmatrix} \begin{matrix} \nwarrow \text{row } i_1 \\ \swarrow \text{row } i_2 \end{matrix}.$$

Theorem 2.11. $\det A^{\text{tr}} = \det A$.

Theorem 2.10. *Let A and B be n by n matrices. Then*

$$\det(A \cdot B) = (\det A) \cdot (\det B).$$

Quick Review of some Linear Algebra

Let V be a vector space. A set $\mathbf{a}_1, \ldots, \mathbf{a}_m$ of vectors in V is said to **span** V if to each $\mathbf{x}$ in V, there corresponds *at least* one m-tuple of scalars $c_1, \ldots, c_m$ such that

$$\mathbf{x} = c_1\mathbf{a}_1 + \cdots + c_m\mathbf{a}_m.$$

In this case, we say that $\mathbf{x}$ can be written as a **linear combination** of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_m$.

Theorem 1.2. *Let V be a vector space of dimension m. If W is a linear subspace of V (different from V), then W has dimension less than m. Furthermore, any basis $\mathbf{a}_1, \ldots, \mathbf{a}_k$ for W may be extended to a basis $\mathbf{a}_1, \ldots, \mathbf{a}_k, \mathbf{a}_{k+1}, \ldots, \mathbf{a}_m$ for V.* □

Theorem 1.6. *If B is the matrix obtained by applying an elementary row operation to A, then*

$$\text{rank } B = \text{rank } A. \quad \square$$

If V has a basis consisting of m vectors, we say that m is the **dimension** of V. We make the convention that the vector space consisting of the zero vector alone has dimension zero.

The set of matrices has, however, an additional operation, called **matrix multiplication.** If A is a matrix of size n by m, and if B is a matrix of size m by p, then the **product** $A \cdot B$ is defined to be the matrix C of size n by p whose general entry c_{ij} is given by the equation

$$c_{ij} = \sum_{k=1}^{m} a_{ik}b_{kj}.$$

The set $\mathbf{a}_1, \ldots, \mathbf{a}_m$ of vectors is said to be **independent** if to each $\mathbf{x}$ in V there corresponds *at most* one m-tuple of scalars $c_1, \ldots, c_m$ such that

$$\mathbf{x} = c_1\mathbf{a}_1 + \cdots + c_m\mathbf{a}_m.$$

Equivalently, $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ is independent if to the zero vector $\mathbf{0}$ there corresponds only one m-tuple of scalars $d_1, \ldots, d_m$ such that

$$\mathbf{0} = d_1\mathbf{a}_1 + \cdots + d_m\mathbf{a}_m,$$

namely the scalars $d_1 = d_2 = \cdots = d_m = 0$.

Theorem 1.4. *Let V be a vector space with basis $\mathbf{a}_1, \ldots, \mathbf{a}_m$. Let W be a vector space. Given any m vectors $\mathbf{b}_1, \ldots, \mathbf{b}_m$ in W, there is exactly one linear transformation $T : V \to W$ such that, for all i, $T(\mathbf{a}_i) = \mathbf{b}_i$.* □

Theorem 1.1. *Suppose V has a basis consisting of m vectors. Then any set of vectors that spans V has at least m vectors, and any set of vectors of V that is independent has at most m vectors. In particular, any basis for V has exactly m vectors.* □

Theorem 1.5. *For any matrix A, the row rank of A equals the column rank of A.* □

(5) For each k, there is a k by k matrix I_k such that if A is any n by m matrix,

$$I_n \cdot A = A \quad \text{and} \quad A \cdot I_m = A.$$

(1) Exchange rows i_1 and i_2 of A (where $i_1 \neq i_2$).
(2) Replace row i_1 of A by itself plus the scalar c times row i_2 (where $i_1 \neq i_2$).
(3) Multiply row i of A by the non-zero scalar λ.

Theorem 1.3. *If A has size n by m, and B has size m by p, then*

$$|A \cdot B| \leq m|A|\,|B|. \quad \square$$

(1) $A \cdot (B \cdot C) = (A \cdot B) \cdot C$.
(2) $A \cdot (B + C) = A \cdot B + A \cdot C$.
(3) $(A + B) \cdot C = A \cdot C + B \cdot C$.
(4) $(cA) \cdot B = c(A \cdot B) = A \cdot (cB)$.

$\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n)$,
$c\mathbf{x} = (cx_1, \ldots, cx_n)$.

(1) $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$.
(2) $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$.
(3) $\langle c\mathbf{x}, \mathbf{y} \rangle = c\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, c\mathbf{y} \rangle$.
(4) $\langle \mathbf{x}, \mathbf{x} \rangle > 0$ if $\mathbf{x} \neq \mathbf{0}$.

(1) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.
(2) $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$.
(3) There is a unique vector $\mathbf{0}$
(4) $\mathbf{x} + (-1)\mathbf{x} = \mathbf{0}$.
(5) $1\mathbf{x} = \mathbf{x}$.
(6) $c(d\mathbf{x}) = (cd)\mathbf{x}$.
(7) $(c + d)\mathbf{x} = c\mathbf{x} + d\mathbf{x}$.
(8) $c(\mathbf{x} + \mathbf{y}) = c\mathbf{x} + c\mathbf{y}$.

$$B = \begin{bmatrix} \circledast & * & * & * & * & * \\ 0 & \circledast & * & * & * & * \\ 0 & 0 & 0 & \circledast & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

(1) $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$.
(2) $T(c\mathbf{x}) = cT(\mathbf{x})$.

$$C = \begin{bmatrix} 1 & 0 & * & 0 & * & * \\ 0 & 1 & * & 0 & * & * \\ 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

$\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 + \cdots + x_ny_n$.

(1) $(A^{\text{tr}})^{\text{tr}} = A$.
(2) $(A + B)^{\text{tr}} = A^{\text{tr}} + B^{\text{tr}}$.
(3) $(A \cdot C)^{\text{tr}} = C^{\text{tr}} \cdot A^{\text{tr}}$.
(4) $\text{rank } A^{\text{tr}} = \text{rank } A$.

(1) $\|\mathbf{x}\| > 0$ if $\mathbf{x} \neq \mathbf{0}$.
(2) $\|c\mathbf{x}\| = |c|\,\|\mathbf{x}\|$.
(3) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

(3′) $\|\mathbf{x} - \mathbf{y}\| \geq \|\mathbf{x}\| - \|\mathbf{y}\|$.

$|\mathbf{x}| = \max\{|x_1|, \ldots, |x_n|\}$.

$T(\mathbf{x}) = A \cdot \mathbf{x}$.

standard basis for $\mathbf{R}^n$:

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0),$$
$$\mathbf{e}_2 = (0, 1, 0, \ldots, 0),$$
$$\cdots$$
$$\mathbf{e}_n = (0, 0, 0, \ldots, 1).$$

$|\mathbf{x}| \leq \|\mathbf{x}\| \leq \sqrt{n}|\mathbf{x}|$.

$\|\mathbf{x}\| = \langle \mathbf{x}, \mathbf{y} \rangle^{1/2}$.