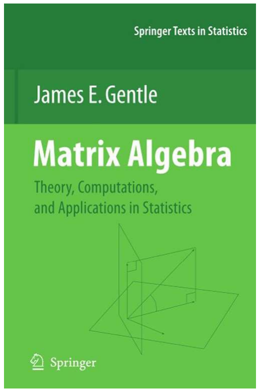
Theory of Matrices

Chapter 3. Basic Properties of Matrices

3.2. Multiplication of Matrices and Multiplication of Vectors and Matrices—Proofs of Theorems



Theorem 3.2.1

Theorem 3.2.1. Properties of Matrix Multiplication.

(1) Let $A = [a_{ij}]$ be $m \times n$ and let $B = [b_{ij}]$ be $n \times s$. Then $(AB)^T = B^T A^T$.

(2) Let $A = [a_{ij}]$ be $m \times n$, $B = [b_{ij}]$ be $n \times s$, and $C = [c_{ij}]$ be $s \times t$. Then $A(BC) = (AB)C$. That is, matrix multiplication is associative.

(3) Let $A = [a_{ij}]$ be $m \times n$ and let $B = [b_{ij}]$ and $C = [c_{ij}]$ be $n \times s$. Then $A(B + C) = AB + AC$. Let $A = [a_{ij}]$ be $m \times n$ and let $B = [b_{ij}]$ and $C = [c_{ij}]$ be $n \times m$ matrices. Then $(B + C)A = BA + CA$. That is, matrix multiplication distributes over matrx addition.

(4) Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be $n \times n$ matrices. If A and B are diagonal then AB is diagonal. If A and B are upper triangular then AB is upper triangular. If A and B are lower triangular then AB is lower triangular.

Theorem 3.2.1 (continued 1)

(1) Let $A = [a_{ij}]$ be $m \times n$ and let $B = [b_{ij}]$ be $n \times s$. Then $(AB)^T = B^T A^T$.

Proof. Let $C = [c_{ij}] = (AB)^T$. The (i,j)th entry of AB is $\sum_{k=1}^{n} a_{ik} b_{kj}$, so $c_{ij} = \sum_{k=1}^{n} a_{jk} b_{ki}$. Let $B^T = [b_{ij}]^T = [b_{ij}^t] = [b_{ji}]$ and $A^T = [a_{ij}]^T = [a_{ij}^t] = [a_{ji}]$. Then the (i,j)th entry of $B^T A^T$ is

$$\sum_{k=1}^{n} b_{ik}^t a_{kj}^t = \sum_{k=1}^{n} b_{ki} a_{jk} = \sum_{k=1}^{n} a_{jk} b_{ki} = c_{ij}$$

and therefore $C = (AB)^T = B^T A^T$. □

Theorem 3.2.1 (continued 2)

(2) Let $A = [a_{ij}]$ be $m \times n$, $B = [b_{ij}]$ be $n \times s$, and $C = [c_{ij}]$ be $s \times t$. Then $A(BC) = (AB)C$. That is, matrix multiplication is associative.

Proof. The (i,j)th entry of BC is $\sum_{k=1}^{s} b_{ik} c_{kj}$ and so the (k,j)th entry of BC is $\sum_{\ell=1}^{s} b_{k\ell} c_{\ell j}$. Therefore the (i,j)th entry of $A(BC)$ is

$$\sum_{k=1}^{n} a_{ik} \left(\sum_{\ell=1}^{s} b_{k\ell} c_{\ell j} \right) = \sum_{\ell=1}^{s} \left(\sum_{k=1}^{n} a_{ik} b_{k\ell} \right) c_{\ell j} = \sum_{k=1}^{s} \left(\sum_{\ell=1}^{n} a_{i\ell} b_{\ell k} \right) c_{kj}$$

where the second inequality holds by interchanging dummy variables ℓ and k. Now $\sum_{\ell=1}^{n} a_{i\ell} b_{\ell k}$ is the (i,k)th entry of AB, and so the last sum is the (i,j)th entry of $(AB)C$. Therefore $A(BC) = (AB)C$. □

Theorem 3.2.1 (continued 3)

(3) Let $A = [a_{ij}]$ be $m \times n$ and let $B = [b_{ij}]$ and $C = [c_{ij}]$ be $n \times s$. Then $A(B + C) = AB + AC$. Let $A = [a_{ij}]$ be $m \times n$ and let $B = [b_{ij}]$ and $C = [c_{ij}]$ be $s \times m$ matrices. Then $(B + C)A = BA + CA$. That is, matrix multiplication distributes over vector addition.

Proof. (3) The (k, j)th entry of $B + C$ is $b_{kj} + c_{kj}$ and so the (i, j)th entry of $A(B + C)$ is

$$\sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj}) = \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj},$$

which is the (i, j)th entry of $AB + AC$, and so $A(B + C) = AB + AC$. Similarly, $(B + C)A = BA + CA$. □

Theorem 3.2.1 (continued 4)

(4) Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be $n \times n$ matrices. If A and B are diagonal then AB is diagonal. If A and B are upper triangular then AB is upper triangular. If A and B are lower triangular then AB is lower triangular.

Proof. (4) The proof is left as Exercise 3.2. □

Theorem 3.2.2

Theorem 3.2.2. Consider partitioned matrices $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ and $\begin{bmatrix} E & F \\ G & H \end{bmatrix}$ where $A = [a_{ij}]$ is $k \times \ell$, $B = [b_{ij}]$ is $k \times m$, $C = [c_{ij}]$ is $n \times \ell$, $D = [d_{ij}]$ is $n \times m$, $E = [e_{ij}]$ is $\ell \times p$, $F = [f_{ij}]$ is $\ell \times q$, $G = [g_{ij}]$ is $m \times p$, and $H = [h_{ij}]$ is $m \times q$. Then the product of the partitioned matrices is the partitioned matrix

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} E & F \\ G & H \end{bmatrix} = \begin{bmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{bmatrix}.$$

Notice that the dimensions of the matrices insure that all matrix products involve matrices conformable for multiplication.

Theorem 3.2.2 (continued 1)

Proof. The dimensions of the matrix products are:

$$\underbrace{\begin{bmatrix} A & B \\ C & D \end{bmatrix}}_{(k+n)\times(\ell+m)} \underbrace{\begin{bmatrix} E & F \\ G & H \end{bmatrix}}_{(\ell+m)\times(p+q)} \begin{bmatrix} \overset{k\times p}{\downarrow} & \overset{k\times q}{\downarrow} \\ AE + BG & AF + BH \\ CE + DG & CF + DH \\ \underset{n\times p}{\uparrow} & \underset{n\times q}{\uparrow} \end{bmatrix}$$

Let $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = [m_{ij}]$ and $N = \begin{bmatrix} E & F \\ G & H \end{bmatrix} = [n_{ij}]$. Then the (i, j)th entry of MN is $\sum_{r=1}^{\ell+m} m_{ir}n_{rj}$. For $i \in \{1, 2, \ldots, k\}$ and $j \in \{1, 2, \ldots, p\}$ we have (where we set $s = r - \ell$) that

$$\sum_{r=1}^{\ell+m} m_{ir}n_{rj} = \sum_{r=1}^{\ell} m_{ir}n_{rj} + \sum_{r=\ell+1}^{\ell+m} m_{ir}n_{rj} = \sum_{r=1}^{\ell} a_{ir}e_{rj} + \sum_{s=1}^{m} b_{is}g_{sj}$$

since $m_{ir} = a_{ir}$ for $r \in \{1, 2, \ldots, \ell\}$, $m_{ir} = b_{ir}$ for $r \in \{\ell + 1, \ell + 2, \ldots, \ell + m\}$, $n_{rj} = e_{rj}$ for $r \in \{1, 2, \ldots, \ell\}$, and ...

Theorem 3.2.2 (continued 2)

Proof. $\ldots n_{rj} = g_{sj}$ for $\{r \in \ell+1, \ell+2, \ldots, \ell+m\}$ (that is, $s \in \{1, 2, \ldots, n\}$) where $i \in \{1, 2, \ldots, k\}$ and $j \in \{1, 2, \ldots, p\}$. Therefore the (i, j)th entry of MN is the sum of the (i, j)th entry of AE and the (i, j)th entry of BG, as claimed.

The result similarly holds for $i \in \{1, 2, \ldots, k\}$ and $j \in \{p+1, p+2, \ldots, p+q\}$ (where the (i, j)th entry of MN is the $(i, j-p)$th entry of $AF + BH$), for $i \in \{k+1, k+2, \ldots, k+n\}$ and $j \in \{1, 2, \ldots, p\}$ (where the (i, j)th entry of MN is the $(i-k, j)$th entry of $CE + DG$), and for $i \in \{k+1, k+2, \ldots, k+n\}$ and $j \in \{p+1, p+2, \ldots, p+q\}$ (where the (i, j)th entry of MN is the $(i-k, j-p)$th entry of $CF + DH$). □

Theorem 3.2.3

Theorem 3.2.3. Each of the three elementary row operations on $n \times m$ matrix A can be accomplished by multiplication on the left by an elementary matrix which is formed by performing the same elementary row operation on the $n \times n$ identity matrix. Each of the three elementary column operations on $n \times m$ matrix A can be accomplished by multiplication on the right by an elementary matrix which is formed by performing the same elementary column operation on the $m \times m$ identity matrix.

Proof. Let $A = [a_{ij}]$ be $n \times m$.
First, consider row interchange. $R_p \leftrightarrow R_q$. Form elementary matrix E_{pq} by interchanging the pth row and qth row of $n \times n$ identity matrix I_n: $I_n \underset{}{\overset{R_p \leftrightarrow R_q}{\frown}} E_{pq} = [e_{ij}]$. So we have $e_{ij} = 0$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$ and $i \neq j$, and $e_{ii} = 1$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$. Now for $i = p$ we have $e_{pj} = 0$ for $j \neq q$, and $e_{pq} = 1$.

Theorem 3.2.3 (continued 1)

Proof (continued). For $i = q$ we have $e_{qj} = 0$ for $j \neq p$, and $e_{qp} = 1$. Let $B = E_{pq}A = [b_{ij}]$. Then B is $n \times m$ and for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$ we have $b_{ij} = \sum_{k=1}^{n} e_{ik} a_{kj} = a_{ij}$, so for these values of i, row i of B is the same as row i of A. For $i = p$, we have $b_{pj} = \sum_{k=1}^{n} e_{pk} a_{kj} = a_{qj}$, so the pth row of B is the same as the qth row of A. For $i = q$, we have $b_{qj} = \sum_{k=1}^{n} e_{qk} a_{kj} = a_{pj}$, so the qth row of B is the same as the pth row of A. That is, $A \overset{R_p \leftrightarrow R_q}{\frown} E_{pq}A$, as claimed.

Second, consider row scaling, $R_p \to sR_p$ where $s \neq 0$. Form elementary matrix E_{sp} by multiplying the pth row of $n \times n$ identity matrix I_n by nonzero scalar s: $I_n \overset{R_p \to sR_p}{\frown} E_{sp} = [e_{ij}]$. So we have $e_{ij} = 0$ for $i \in \{1, 2, \ldots, n\} \setminus \{p\}$ and $i \neq j$, and $e_{ii} = 1$ for $i \in \{1, 2, \ldots, n\} \setminus \{p\}$. Now for $i = p$ we have $e_{pj} = 0$ for $j \neq p$ and $e_{pp} = s$. Let $B = E_{sp}A = [b_{ij}]$.

Theorem 3.2.3 (continued 2)

Proof (continued). Then B is $n \times m$ and for $i \in \{1, 2, \ldots, n\} \setminus \{p\}$, $b_{ij} = \sum_{k=1}^{n} e_{ik} a_{kj} = a_{ij}$, so for these values of i, the ith row of B is the same as the ith row of A. For $i = p$, $b_{pj} = \sum_{k=1}^{n} e_{pk} a_{kj} = sa_{pj}$, so the pth row of B is s times the pth row of A. That is, $A \overset{R_p \to sR_p}{\frown} E_{sp}A$, as claimed.

Third, consider row addition: $R_p \to R_p + sR_q$. We leave this as Exercise 3.D.

The proof for elementary column operations is similar. □

Theorem 3.2.4

Theorem 3.2.4. For $n \times n$ matrices A and B, $\det(AB) = \det(A)\det(B)$.

Proof. We have

$$\det(A)\det(B) = \det\left(\begin{bmatrix} A & 0 \\ -I_n & B \end{bmatrix}\right) \text{ by Theorem 3.1.G}$$

$$= \det\left(\begin{bmatrix} A & 0 \\ -I_n & B \end{bmatrix}\begin{bmatrix} I_n & B \\ 0 & I_n \end{bmatrix}\right) \text{ by Theorem 3.1.H since}$$

$$\left(\begin{bmatrix} I_n & B \\ 0 & I_n \end{bmatrix}\right) \text{ is upper triangular with diagonal entries 1}$$

$$= \det\left(\begin{bmatrix} A & AB \\ -I_n & 0 \end{bmatrix}\right) \text{ by Theorem 3.2.2.}$$

Now if we swap rows i and $n+i$ for $i = 1, 2, \ldots, n$ of $\begin{bmatrix} A & AB \\ -I_n & 0 \end{bmatrix}$ then we get $\begin{bmatrix} -I_n & 0 \\ A & AB \end{bmatrix}$ and by Theorem 3.1.C, . . .

Theorem 3.2.4 (continued)

Theorem 3.2.4. For $n \times n$ matrices A and B, $\det(AB) = \det(A)\det(B)$.

Proof (continued). . . .

$$(-1)^n \det\left(\begin{bmatrix} -I_n & 0 \\ A & AB \end{bmatrix}\right) = (-1)^n \det(-I_n)\det(AB) = (-1)^{2n}\det(AB)$$

since $\det(-I_n) = (-1)^n$ by Example 3.1.A. Therefore,

$$\det(A)\det(B) = \det\left(\begin{bmatrix} A & AB \\ -I_n & 0 \end{bmatrix}\right)$$

$$= (-1)^n \det\left(\begin{bmatrix} -I_n & 0 \\ A & AB \end{bmatrix}\right) = (-1)^{2n}\det(AB) = \det(AB).$$

□

Theorem 3.2.7

Theorem 3.2.7. Let A and B be $n \times n$ matrices. Then $\mathrm{tr}(A \otimes B) = \mathrm{tr}(A)\mathrm{tr}(B)$.

Proof. From $(*)$, the diagonal entries of $A \otimes B$ are $a_{1+\lfloor(i-1)/n\rfloor, 1+\lfloor(i-1)/n\rfloor} b_{i-n\lfloor(i-1)/n\rfloor, i-n\lfloor(i-1)/n\rfloor}$ for $i \in \{1, 2, \ldots, n^2\}$. For $i \in \{1+(k-1)n, 2+(k-1)n, \ldots, kn\}$ we have $1 + \lfloor(i-1)/n\rfloor = 1 + (k-1) = k$ and $i - n\lfloor(i-1)/n\rfloor = i - n(k-1) \in \{1, 2, \ldots, n\}$, and for these values of i, $a_{1+\lfloor(i-1)/n\rfloor, 1+\lfloor(i-1)/n\rfloor} b_{i-n\lfloor(i-1)/n\rfloor, i-n\lfloor(i-1)/n\rfloor} = a_{kk} b_{i-n(k-1), i-n(k-1)}$. So

$$\mathrm{tr}(A \otimes B) = \sum_{i=1}^{n^2} a_{1+\lfloor(i-1)/n\rfloor, 1+\lfloor(i-1)/n\rfloor} b_{i-n\lfloor(i-1)/n\rfloor, i-n\lfloor(i-1)/n\rfloor}$$

$$= \sum_{k=1}^{n} \sum_{i=1+(k-1)n}^{kn} a_{kk} b_{i-n(k-1), i-n(k-1)} = \sum_{k=1}^{n} a_{kk} \sum_{i=1}^{n} b_{ii} = \mathrm{tr}(A)\mathrm{tr}(B).$$

□

Theorem 3.2.8

Theorem 3.2.8. Properties of the Inner Product of Matrices. Let A, B, and C be matrices conformable for the addition and inner products given below. Then

(1) If $A \neq 0$ then $\langle A, A\rangle > 0$ and $\langle 0, A\rangle = \langle A, 0\rangle = \langle 0, 0\rangle = 0$.
(2) $\langle A, B\rangle = \langle B, A\rangle$.
(3) $\langle sA, B\rangle = s\langle A, B\rangle = \langle A, sB\rangle$ for scalar $s \in \mathbb{R}$.
(4) $\langle (A+B), C\rangle = \langle A, C\rangle + \langle B, C\rangle$ and $\langle C, (A+B)\rangle = \langle C, A\rangle + \langle C, B\rangle$.
(5) $\langle A, B\rangle = \mathrm{tr}(A^T B)$.
(6) $\langle A, B\rangle = \langle A^T, B^T\rangle$.
(7) Schwarz Inequality: For $n \times m$ matrices A and B, $|\langle A, B\rangle| = \langle A, A\rangle^{1/2}\langle B, B\rangle^{1/2}$.

Proof. We leave the proofs of (5) and (7) as an exercise.

Theorem 3.2.8 (continued 1)

(1) If $A \neq 0$ then $\langle A, A\rangle > 0$ and $\langle 0, A\rangle = \langle A, 0\rangle = \langle 0, 0\rangle = 0$.

Proof. (1) Let the column vectors of $n \times m$ matrix A be $a_1, a_2, \ldots, a_m$. If $A \neq 0$ then for some $1 \leq k \leq m$ we have $\|a_k\| > 0$. So

$$\langle A, A\rangle = \sum_{j=1}^{m} a_j^T a_j = \sum_{j=1}^{m} \langle a_j, a_j\rangle = \sum_{j=1}^{m} \|a_j\|^2 \geq \|a_k\|^2 > 0.$$

The columns of the $n \times m$ zero matrix are all 0 vectors, say $0_1, 0_2, \ldots, 0_m$, and so

$$\langle A, 0\rangle = \sum_{j=1}^{m} a_j^T 0_j = 0 = \sum_{j=1}^{m} 0_j^T a_j = \langle 0, A\rangle,$$

and $\langle 0, 0\rangle = \sum_{j=1}^{m} 0_j^T 0_j = 0$. □

Theorem 3.2.8. Properties of the Inner Product of Matrices

Theorem 3.2.8 (continued 2)

(2) $\langle A, B\rangle = \langle B, A\rangle$.
(3) $\langle sA, B\rangle = s\langle A, B\rangle = \langle A, sB\rangle$ for scalar $s \in \mathbb{R}$.

Proof. (2) Let the column vectors of $n \times m$ matrices A and B be $a_1, a_2, \ldots, a_m$ and $b_1, b_2, \ldots, b_m$, respectively. Then

$$\begin{aligned}\langle A, B\rangle &= \sum_{j=1}^{m} a_j^T b_j = \sum_{j=1}^{m} \langle a_j, b_j\rangle \\ &= \sum_{j=1}^{m} \langle b_j, a_j\rangle \text{ by Theorem 2.1.6(2)} \\ &= \sum_{j=1}^{m} b_j^T a_j = \langle B, A\rangle. \quad \square\end{aligned}$$

(3) Let the column vectors of $n \times m$ matrix A be $a_1, a_2, \ldots, a_m$ and let $s \in \mathbb{R}$ be a scalar. Then the column vectors of sA are $sa_1, sa_2, \ldots, sa_m$ and so ...

Theorem 3.2.8. Properties of the Inner Product of Matrices

Theorem 3.2.8 (continued 3)

(3) $\langle sA, B\rangle = s\langle A, B\rangle = \langle A, sB\rangle$ for scalar $s \in \mathbb{R}$.

Proof (continued). ...

$$\langle sA, B\rangle = \sum_{j=1}^{m} (sa_j)^T b_j = \sum_{j=1}^{m} sa_j^T b_j = s\sum_{j=1}^{m} a_j^T b_j = s\langle A, B\rangle.$$

Also

$$\begin{aligned}\langle A, sB\rangle &= \langle sB, A\rangle \text{ by part (2)} \\ &= s\langle B, A\rangle \text{ as just shown} \\ &= s\langle A, B\rangle \text{ by part (2).} \quad \square\end{aligned}$$

Theorem 3.2.8. Properties of the Inner Product of Matrices

Theorem 3.2.8 (continued 4)

(4) $\langle (A + B), C\rangle = \langle A, C\rangle + \langle B, C\rangle$ and $\langle C, (A + B)\rangle = \langle C, A\rangle + \langle C, B\rangle$.

Proof. (4) Let A, B, C be $n \times m$ matrices with column vectors $a_1, a_2, \ldots, a_m$, $b_1, b_2, \ldots, b_m$, and $c_1, c_2, \ldots, c_m$, respectively. Then the columns of $A + B$ are $a_1 + b_1, a_2 + b_2, \ldots, a_m + b_m$ and

$$\langle (A + B), C\rangle = \sum_{j=1}^{m} (a_j + b_j)^T c_j = \sum_{j=1}^{m} (a_j^T + b_j^T)c_j = \sum_{j=1}^{m} (a_j^T c_j + b_j^T c_j)$$

$$= \sum_{j=1}^{m} a_j^T c_j + \sum_{j=1}^{m} b_j^T c_j = \langle A, C\rangle + \langle B, C\rangle.$$

Next,

$$\begin{aligned}\langle C, (A + B)\rangle &= \langle (A + B), C\rangle \text{ by part (2)} \\ &= \langle A, C\rangle + \langle B, C\rangle \text{ as just shown} \\ &= \langle C, A\rangle + \langle C, B\rangle \text{ by part (2).} \quad \square\end{aligned}$$

Theorem 3.2.8 (continued 5)

(6) $\langle A, B\rangle = \langle A^T, B^T\rangle$.

Proof. (6) Let $A = [a_{ij}]$, $A^T = [a^t_{ij}]$, $B = [b_{ij}]$, and $B^T = [b^t_{ij}]$ be $n \times m$ matrices (so $a^t_{ij} = a_{ji}$ and $b^t_{ij} = b_{ji}$) and denote the columns of A as a_j, the columns of B as b_j, the columns of A^T as a^t_j, and the columns of B^T as b^t_j. Then

$$\langle A, B\rangle = \sum_{j=1}^{m} a_j^T b_j = \sum_{j=1}^{m}\left(\sum_{i=1}^{n} a_{ij} b_{ij}\right)$$

$$= \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ji} b_{ji}\right) \text{ interchanging } i \text{ and } j$$

$$= \sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ji} b_{ji}\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} a^t_{ij} b^t_{ij}\right) = \sum_{j=1}^{n} (a^t_j)^T b^t_j = \langle A^T, B^T\rangle.$$

□