## Chapter 3. Basic Properties of Matrices

Note. This long chapter (over 100 pages) contains the bulk of the material for this course. As in Chapter 2, unless stated otherwise, all matrix entries are real numbers. Chapters 8 and 9 include material on applications of this material to regression and statistics.

## Section 3.1. Basic Definitions and Notation

Note. In this section, we reintroduce many of the definitions you should be familiar with from sophomore Linear Algebra (diagonal, trace, minor matrix, cofactor, determinant). We also introduce some notation and matrix/vector manipulations which might be new to you.

Definition. Let $A$ be an $n \times m$ matrix. Treating the columns of $A$ as $n$-vectors, the vector space generated by these column vectors (that is, the span of the columns of $A$ ) is the column space of matrix $A$. This is sometimes called the range or manifold of $A$ and is denoted $\operatorname{span}(A)$. Treating the rows of $A$ as $m$-vectors, the vectors generated by these row vectors (that is, the span of the rows of $A$ ) is the row space of matrix $A$.

Definition. For matrix $A=\left[a_{i j}\right]$, define scalar multiplication entry-wise as $c A=$ $c\left[a_{i j}\right]=\left[c a_{i j}\right]$. For $A=\left[a_{i j}\right]$ and $B=\left[b_{i j}\right] m \times n$ matrices, define matrix addition entry-wise as $A+B=\left[a_{i j}\right]+\left[b_{i j}\right]=\left[a_{i j}+b_{i j}\right]$. The diagonal elements of $A=$ $\left[a_{i j}\right]$ are the elements $a_{i i}$ for $i=1,2, \ldots, \min \{m, n\}$ (these elements make up the principal diagonal), elements $a_{i j}$ where $i<j$ are above the diagonal, and elements $a_{i j}$ where $i>j$ are below the diagonal. The elements $a_{i, i+k}$ (for fixed $k$ ) form the $k$ th codiagonal or $k$ th minor diagonal. For $n \times m$ matrix $A$, the elements $a_{i, m+1-i}$ for $i=1,2, \ldots, \min \{m, n\}$ are the skew diagonal elements of $A$.

Definition. Square matrix $A$ of size $n \times n$ is symmetric if $a_{i j}=a_{j i}$ for all $1 \leq$ $i, j \leq n$. Square matrix $A$ is skew symmetric if $a_{i j}=-a_{j i}$ for $1 \leq i, j \leq n$. Square matrix $A$ with complex entries is Hermetian (or self-adjoint) if $a_{i j}=\overline{a_{j i}}$ (where $\overline{a_{j i}}$ represents the complex conjugate of $a_{j i}$ ).

Definition. If all entries of a square matrix are 0 except for the principal diagonal elements, then matrix $A$ is a diagonal matrix. If all the principal diagonal elements are 0 , the matrix is a hollow matrix. If all entries except the principal skew diagonal elements of a matrix are 0 , the matrix is a skew diagonal matrix.

Note. A skew symmetric matrix must satisfy $a_{i i}=-a_{i i}$ for $1 \leq i \leq n$ and so is necessarily hollow.

Definition. An $n \times m$ matrix $A$ for which $\left|a_{i i}\right|>\sum_{j=1, j \neq i}^{m}\left|a_{i j}\right|$ for each $i=$ $1,2, \ldots, n$ is row diagonally dominant (that is, the absolute value of the diagonal entry is larger than the sum of the absolute values of the other entries in the row containing that diagonal entry [for each row]). Similarly, if $\left|a_{j j}\right|>\sum_{i=1, i \neq j}^{n}\left|a_{i j}\right|$ for each $j=1,2, \ldots, m$ then $A$ is column diagonally dominant.

Note. If $A$ is symmetric (and hence square) then row and column diagonal dominance are equivalent in which case $A$ is simply called diagonally dominant.

Definition. A square matrix such that all elements below the diagonal are 0 is an upper triangular matrix; if all entries above the diagonal are 0 the matrix is a lower triangular matrix. If a nonsquare matrix has all entries below the diagonal $\left\{a_{i i}\right\}$ equal to 0 (or all entries above the diagonal equal to 0 ) then the matrix is trapezoidal.

Definition. A square matrix $A$ for which all elements are 0 except for $a_{i, i+c_{k}}$ for some ("small") $c_{k}$ where $c_{k} \in\left\{-w_{\ell},-w_{\ell}+1, \ldots,-1,0,1, \ldots, w_{u-1}, w_{u}\right\}$ is a band matrix (or banded matrix). Value $w_{\ell}$ is the lower band width and $w_{u}$ is the upper band width. A band matrix with lower and upper band widths of 1 is tridiagonal. A matrix is in upper Hessenburg form if it is upper triangular except for the first (lower) subdiagonal.

Note. Gentle states that band matrices arise in time series, stochastic processes, and differential equations (see page 43).

Definition. For $m \times n$ matrix $A=\left[a_{i j}\right]$, the transpose matrix is the $n \times m$ matrix $A^{T}=\left[a_{j i}\right]$. If the entries of $A$ are complex numbers, then the adjoint of $A$ is the conjugate transpose $A^{H}=\left[\overline{a_{j i}}\right]$, also called the Hermetian of $A$.

Note. As you've seen in sophomore Linear Algebra, we often put vectors into matrices and extract vectors from matrices. We now define such operations formally.

Definition. Define the "constructor function" diag : $\mathbb{R}^{n} \rightarrow \mathbb{R}^{n \times n}$ as

$$
\operatorname{diag}\left(\left[d_{1}, d_{2}, \ldots, d_{n}\right]\right)=\left[\begin{array}{ccccc}
d_{1} & 0 & 0 & \cdots & 0 \\
0 & d_{2} & 0 & \cdots & 0 \\
0 & 0 & d_{3} & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & d_{n}
\end{array}\right]
$$

which maps an $n$-vector into an $n \times n$ diagonal matrix with diagonal entries equal to the entries in the $n$-vector. Define vecdiag : $\mathbb{R}^{n \times m} \rightarrow \mathbb{R}^{k}$ where $k=\min \{m, n\}$ as

$$
\operatorname{vecdiag}(A)=\left[a_{11}, a_{22}, \ldots, a_{k k}\right] .
$$

Define vec : $\mathbb{R}^{n \times m} \rightarrow \mathbb{R}^{n m}$ as

$$
\operatorname{vec}(A)=\left[a_{1}^{T}, a_{2}^{T}, \ldots, a_{m}^{T}\right]
$$

where $a_{1}, a_{2}, \ldots, a_{m}$ are the columns of matrix $A$ (so " $\left[a_{1}^{T}, a_{2}^{T}, \ldots, a_{m}^{T}\right]$ " denotes an $n m$ row vector). Define vech $: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m(m+1) / 2}$ as

$$
\operatorname{vech}(A)=\left[a_{11}, a_{21}, \ldots, a_{m 1}, a_{22}, a_{32}, \ldots, a_{m 2}, a_{33}, a_{43}, \ldots, a_{m m}\right]
$$

(Gentle restricts vech to the symmetric matrices in $\mathbb{R}^{m \times n}$ so that vech $(A)$ produces a vector with entries as those entries on and above the diagonal of $A$ ).

Note. Gentle doesn't give a formal definition of a partitioned matrix. You are familiar with the idea of a partitioned matrix since the system of equations $A \vec{x}=$ $\vec{b}$ (in the notation of sophomore Linear Algebra) is dealt with in terms of the partitioned matrix (or "augmented matrix") $[A \mid \vec{b}]$.
"Definition." An $n \times m$ matrix $A$ can be partitioned into four matrices $A_{11}, A_{12}, A_{21}$, and $A_{22}$ as

$$
A=\left[\begin{array}{ll}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{array}\right]
$$

where $A_{11}$ and $A_{12}$ have the same number of rows (say $r_{1}$ ); $A_{21}$ and $A_{22}$ have the same number of rows (say $r_{2}$ ); $A_{11}$ and $A_{21}$ have the same number of columns (say $c_{1}$ ); and $A_{12}$ and $A_{22}$ have the same number of columns (say $c_{2}$ ). The $i$ th row of $A_{11}$ "combined with" the $i$ th row of $A_{12}$ gives the $i$ th row of $A$ for $1 \leq i \leq r_{1}$; the $i$ th row of $A_{21}$ "combined with" the $i$ th row of $A_{22}$ gives the $\left(r_{1}+i\right)$ th row of $A$ for $1 \leq i \leq r_{2}$; the $j$ th column of $A_{11}$ "combined with" the $j$ th column of $A_{21}$ gives the $j$ th column of $A$ for $1 \leq j \leq c_{1}$; and the $j$ th column of $A_{12}$ "combined with" the $j$ th column of $A_{22}$ gives the $\left(c_{1}+j\right)$ th column of $A$ for $1 \leq j \leq c_{2}$.

Definition. Matrix $B$ is a submatrix of matrix $A$ if $B$ is obtained from $A$ by deleting some of the rows and some of the columns of $A$. A square submatrix whose principal diagonal elements are elements of the principal diagonal of matrix $A$ is a principal submatrix of $A$. A principal submatrix $B$ obtained from $n \times m$ matrix $A$ such that if column $j$ is eliminated from $A$ then all columns $j+1, j+2, \ldots m$ are eliminated in producing $B$, and if row $i$ is eliminated from $A$ then all rows $i+1, i+2, \ldots, n$ are eliminated, is a leading principal submatrix.

Theorem 3.1.1. Suppose matrix $A$ is diagonally dominant (that is, $A$ is symmetric and row and column diagonally dominant). If $B$ is a principal submatrix of $A$ then $B$ is also diagonally dominant.

Definition. A partitioned matrix is block diagonal if it is of the form

$$
\left[\begin{array}{cccc}
A_{1} & 0 & \cdots & 0 \\
0 & A_{2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & A_{k}
\end{array}\right]
$$

for $i=1,2, \ldots, k$ and where " 0 " represents a submatrix of all 0 's of appropriate dimensions to yield a partitioned matrix. Similar to the function diag : $\mathbb{R}^{n} \rightarrow \mathbb{R}^{n \times n}$, we denote the matrix above as $\operatorname{diag}\left(A_{1}, A_{2}, \ldots, A_{k}\right)$.

Definition. Let $A=\left[a_{i j}\right]$ and $B=\left[b_{i j}\right]$ be $n \times m$ matrices. The sum $A+B$ is the $n \times m$ matrix $C=\left[c_{i j}\right]$ where $c_{i j}=a_{i j}+b_{i j}$. For $r \in \mathbb{R}$ a scalar, $r A$ is the matrix $D=\left[d_{i j}\right]$ where $d_{i j}=r a_{i j}$.

Theorem 3.1.2. Let $A$ and $B$ be $n \times m$ matrices and $r, s \in \mathbb{R}$. Then
(a) $A+B=B+A$ (Commutivity Law of Matrix Addition).
(b) $(A+B)+C=A+(B+C)$ (Associative Law of Matrix Addition).
(c) $A+0=0+A$ where 0 is the $n \times m$ zero matrix (Matrix Additive Identity).
(d) $r(A+B)=r A+r B$ (Distribution of Scalar Multiplication over Matrix Addition).
(e) $(r+s) A=r A+s A$ (Distribution of Scalar Multiplication over Scalar Addition). (f) $(r s) A=r(s A)$ (Associativity Law of Scalar Multiplication).

Note. Since we have defined matrix addition and scalar multiplication, we can show that $\mathbb{R}^{n \times m}$ is in fact a vector space (of dimension $n m$ ). A subspace of $\mathbb{R}^{n \times m}$ is the space of all symmetric $n \times n$ matrices; $\mathbb{R}^{n \times n}$ is of dimension $n(n+1) / 2$ (see Exercise 3.1).

Notation. For $a \in \mathbb{R}$ and $A=\left[a_{i j}\right]$ we define $A+a=\left[a_{i j}+a\right]$. If all elements of real matrix $A$ are positive, we write $A>0$; if all elements of $A$ are nonnegative we write $A \geq 0$.

Definition. The trace of $n \times n$ matrix $A=\left[a_{i j}\right]$ is $\operatorname{tr}(A)=\sum_{i=1}^{n} a_{i i}$.

Note. Three properties of the trace are $\operatorname{tr}(A)=\operatorname{tr}\left(A^{T}\right), \operatorname{tr}(c A)=c \operatorname{tr}(A)$, and $\operatorname{tr}(A+B)=\operatorname{tr}(A)+\operatorname{tr}(B)$.

Note. To define the determinant of a matrix, we must step aside and study permutations on the set $\{1,2, \ldots, n\}$.

Definition. A permutation of a set $A$ is a one to one and onto function (or "bijection") $\varphi: A \rightarrow A$.

Note. There are $n$ ! permutations on the set $\{1,2, \ldots, n\}$. We use "cyclic notation" to represent such permutations, so

is represented as $(1)(2,3,4)(5,6)$. We "multiply" permutations by reading from right to left: $(1,6)(1,2)(2,1)(1,5)(1,4)=(1,4,5,6)(2)$.We often do not write cycles of length 1 (such as (2) here). A cycle of length 2 is a "transposition." Every cycle is a product of transpositions: $(1,2, \ldots, n)=(1, n)(1, n-1)(1, n-2) \cdots(1,3)(1,2)$. In fact, every permutation of a finite set of at least two elements is a product of transpositions (see Corollary 9.12 of my online notes II.9. "Orbits, Cycles, and the Alternating Groups"). Another important property of transpositions is: No permutation of a finite set can be expressed as a product of an even number of transpositions and as a product of an odd number of transpositions (see Theorem 9.15 of the online notes mentioned above). This property allows us to define a permutation as even or odd according to whether the permutation can be expressed as a product of an even number of transpositions or the product of an odd number of transpositions. Recall that the alternating group $A_{n}$ is the group of all even permutations on the set $\{1,2, \ldots, n\}$.

Example. Consider $\{1,2\}$. There are two permutations: (1)(2) and (1,2). There are 6 permutations of $\{1,2,3\}:(1)(2)(3),(1)(2,3),(1,3)(2),(1,2)(3),(1,2,3)$, and $(1,3,2)$. We have

## Even

$$
\begin{array}{ll}
(1)(2)(3)=(1,2)(1,2) & (1)(2,3)=(2,3) \\
(1,2,3)=(1,3)(1,2) & (1,3)(2)=(1,3) \\
(1,3,2)=(1,2)(1,3) & (1,2)(3)=(1,2)
\end{array}
$$

Definition. Define $\sigma: S_{n} \rightarrow\{-1,1\}$, where $S_{n}$ denotes the group of all permutations of $\{1,2, \ldots, n\}$, as $\sigma(\pi)=1$ if $\pi \in S_{n}$ is even and $\sigma(\pi)=-1$ if $\pi \in S_{n}$ is odd.

Note. Recall in sophomore Linear Algebra (MATH 2010) we have the following determinants:

$$
\operatorname{det}\left[a_{11}\right]=a_{11}, \quad \operatorname{det}\left[\begin{array}{ll}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{array}\right]=a_{11} a_{22}-a_{12} a_{21},
$$

$\operatorname{det}\left[\begin{array}{lll}a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33}\end{array}\right]=a_{11} a_{22} a_{33}+a_{12} a_{23} a_{31}+a_{13} a_{21} a_{32}-a_{11} a_{23} a_{32}-a_{13} a_{22} a_{31}-a_{12} a_{21} a_{33}$.
When $n=2$, with $\pi_{1}=(1)(2)$ and $\pi_{2}=(1,2)$ we have

$$
\sigma\left(\pi_{1}\right) a_{1 \pi_{1}(1)} a_{2 \pi_{1}(2)}+\sigma\left(\pi_{2}\right) a_{1 \pi_{2}(1)} a_{2 \pi_{2}(2)}=+a_{11} a_{22}-a_{12} a_{21}=\operatorname{det}\left[\begin{array}{ll}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{array}\right] .
$$

When $n=3$, with $\pi_{1}=(1)(2)(3), \pi_{2}=(1,2,3), \pi_{3}=(1,3,2)$ (the even permutations) and $\pi_{4}=(2,3), \pi_{5}=(1,3), \pi_{6}=(1,2)$ (the odd permutations) we have

$$
\begin{aligned}
\sum_{j=1}^{6} \sigma\left(\pi_{j}\right) \prod_{i=1}^{3} a_{i \pi_{j}(i)}= & +a_{11} a_{22} a_{33}+a_{12} a_{23} a_{31}+a_{13} a_{21} a_{32} \\
& -a_{11} a_{23} a_{32}-a_{13} a_{22} a_{31}-a_{12} a_{21} a_{33}=\operatorname{det}\left[\begin{array}{ccc}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & a_{33}
\end{array}\right] .
\end{aligned}
$$

We use this pattern to motivate the following definition.

Definition. Let $A=\left[a_{i j}\right]$ be an $n \times n$ matrix and let $S_{n}$ be the set of all permutations on set $\{1,2, \ldots, n\}$. The determinant of $A$, $\operatorname{denoted} \operatorname{det}(A)=|A|$, is

$$
\operatorname{det}(A)=\sum_{\pi \in S_{n}} \sigma(\pi) \prod_{i=1}^{n} a_{i \pi(i)}
$$

Note. Notice that $\operatorname{det}(A)$ involves all possible products where each factor is obtained by picking one and only one element from each row and column of $A$. This definition and the following "lettered" results are based in part on David Harville's Matrix Algebra from a Statistician's Point of View, Springer-Verlag (1997).

Example 3.1.A. We claim that if $A=\left[a_{i j}\right]$ is an $n \times n$ (upper or lower) triangular matrix then $\operatorname{det}(A)=a_{11} a_{22} \cdots a_{n n}$. Say $A$ is upper triangular. By definition, $\operatorname{det}(A)=\sum_{\pi \in S_{n}} \sigma(\pi) \prod_{i=1}^{n} a_{i \pi(i)}$. If $\pi(1) \neq 1$ then for some $k$ with $2 \leq k \leq n$ we have $\pi(k)=1$. But since $k>1$, then $a_{k \pi(k)}=a_{k 1}=0$ and so $\prod_{i=1}^{n} a_{i \pi(i)}=0$. So the only way for $\prod_{i=1}^{n} a_{i \pi(i)}$ to be nonzero is for $\pi(1)=1$. With $\pi(1)=1$ and $\prod_{i=1}^{n} a_{i \pi(i)}$ nonzero, we can show by induction that $\pi(i)=i$ for all $1 \leq i \leq n$; that is, $\pi$ is the identity permutation (which is even). $\operatorname{So} \operatorname{det}(A)=\prod_{i=1}^{n} a_{i i}$, as claimed.

Note 3.1.A. If $n \times n$ matrix $A$ is upper (or lower) triangular, then $\prod_{i=1}^{n} a_{i \pi(i)}=0$ unless $\pi(i) \geq i$ for each $1 \leq i \leq n$. We can only have $\pi(i) \geq i$ for each $1 \leq i \leq n$ if $\pi(i)=i$ for each $1 \leq i \leq n$ (use backwards induction: $\pi(n)=n$, so $\pi(n-1)=n-1$, $\ldots$..). So the only permutation $\pi$ for which $\prod_{i=1}^{n} a_{i \pi(i)}$ may be nonzero is the identity permutation (which is even). So for such $A, \operatorname{det}(A)=\prod_{i=1}^{n} a_{i i}$. Of course a diagonal matrix is upper triangular and its determinant is also the product of its diagonal elements.

Theorem 3.1.A. Let $A=\left[a_{i j}\right]$ be an $n \times n$ matrix. Then $\operatorname{det}(A)=\operatorname{det}\left(A^{T}\right)$.

Theorem 3.1.B. If an $n \times n$ matrix $B$ is formed from a $n \times n$ matrix $A$ by multiplying all of the elements of one row or one column of $A$ by the same scalar $k$ (and leaving the elements of the other $n-1$ row or columns unchanged) then $\operatorname{det}(B)=k \operatorname{det}(A)$.

Note 3.1.B. If a row (or column) of a matrix $A$ is all 0 's then $\operatorname{det}(A)=0$. This follows from Theorem 3.1.B by letting $B=A$ and $k=0$. With $k \in \mathbb{R}$, $\operatorname{det}(k A)=k^{n} \operatorname{det}(A)$; this follows by applying Theorem 3.1.B $n$ times as each row of $A$ is multiplied by $k$. Notice that $\operatorname{det}(-A)=(-1)^{n} \operatorname{det}(A)$.

Note. Theorem 3.1.B shows how the elementary row operation of multiplying a row (or column) by a scalar affects the determinant of a matrix. We now explore the effect of other elementary row operations.

Theorem 3.1.C. If a $n \times n$ matrix $B=\left[b_{i j}\right]$ is formed from an $n \times n$ matrix $A=\left[a_{i j}\right]$ by interchanging two rows (or columns) of $A$ then $\operatorname{det}(B)=-\operatorname{det}(A)$.

Note 3.1.C. If two rows (or columns) of square matrix $A$ are equal, then Theorem 3.1.C implies that $\operatorname{det}(A)=0$. This can be used to prove the following (we leave the detailed proof as Exercise 3.1.A).

Corollary 3.1.D. If a row or column of $n \times n$ matrix $A$ is a scalar multiple of another row or column (respectively) of $A$, then $\operatorname{det}(A)=0$.

Note. The next result, along with Theorem 3.1.B and Theorem 3.1.C, show how elementary row operations affect the determinant of a matrix.

Theorem 3.1.E. Let $B$ represent a matrix formed from $n \times n$ matrix $A$ by adding to any row (or column) of $A$, scalar multiples of one or more other rows (or columns). Then $\operatorname{det}(B)=\operatorname{det}(A)$.

Note. In sophomore Linear Algebra (MATH 2010), the determinant of a square matrix is defined recursively using cofactors. See my online notes on 4.2. The Determinant of a Square Matrix; notice Definition 4.1 and Theorem 4.2. We now show that our definition is equivalent to that definition.

Definition. Let $A=\left[a_{i j}\right]$ be an $n \times n$ matrix. Let $A_{i j}$ represent the $(n-1) \times(n-1)$ submatrix of $A$ obtained by eliminating the $i$ th row and $j$ th column of $A$. The determinant $\operatorname{det}\left(A_{i j}\right)$ is the minor of element $a_{i j}$. The signed minor $(-1)^{i+j} \operatorname{det}\left(A_{i j}\right)$ is the cofactor of $a_{i j}$, denoted $\alpha_{i j}$.

Theorem 3.1.F. Let $A=\left[a_{i j}\right]$ be an $n \times n$ matrix and let $\alpha_{i j}$ represent the cofactor of $a_{i j}$. Then

$$
\begin{equation*}
\operatorname{det}(A)=\sum_{j=1}^{n} a_{i j} \alpha_{i j} \text { for } i=1,2, \ldots, n \tag{5.1}
\end{equation*}
$$

and

$$
\begin{equation*}
\operatorname{det}(A)=\sum_{i=1}^{n} a_{i j} \alpha_{i j} \text { for } j=1,2, \ldots, n \tag{5.2}
\end{equation*}
$$

Definition. The adjoint of a $n \times n$ matrix $A$ is the transpose of the matrix of cofactors of $A,\left[\alpha_{i j}\right]^{T}$, denoted $\operatorname{adj}(A)$. (Gentle calls this the "adjugate" of $A$.)

Theorem 3.1.3. Let $A$ be an $n \times n$ matrix with $\operatorname{adjoint} \operatorname{adj}(A)=\left[\alpha_{i j}\right]^{T}$. Then $A \operatorname{adj}(A)=\operatorname{adj}(A) A=\operatorname{det}(A) I_{n}$.

Note. Notice that Theorem 3.1.3 implies that $A^{-1}=(1 / \operatorname{det}(A)) \operatorname{adj}(A)$. So for $A=\left[\begin{array}{ll}a & b \\ c & d\end{array}\right]$, where $\operatorname{det}(A)=a d-b c \neq 0$, we have $A^{-1}=\frac{1}{a d-b c}\left[\begin{array}{cc}d & -b \\ -c & a\end{array}\right]$.

Note. We now consider the determinant of a partitioned matrix of a certain form.

Theorem 3.1.G. Let $T$ be an $m \times m$ matrix, $V$ an $n \times m$ matrix, $W$ an $n \times n$ matrix, and let ' 0 ' represent the $m \times n$ matrix of all entries as 0 . Then the determinant of the partitioned matrix is

$$
\operatorname{det}\left(\left[\begin{array}{cc}
T & 0 \\
V & W
\end{array}\right]\right)=\operatorname{det}\left(\left[\begin{array}{cc}
W & V \\
0 & T
\end{array}\right]\right)=\operatorname{det}(T) \operatorname{det}(W)
$$

Note. The following will be useful in showing that the determinant of a product of matrices in the next section. This is Harville's Theorem 13.2.11.

Theorem 3.1.H. Let $A$ be $n \times n$ and let $T$ be an $n \times n$ upper or lower triangular matrix with entries of 1 along the diagonal. Then $\operatorname{det}(A T)=\operatorname{det}(T A)=\operatorname{det}(A)$.

Note. The actual computation of the determinant of a large matrix can be extremely time consuming. An efficient way to compute a determinant is to use elementary row operations to reduce the matrix and to track how the reduction effects the determinant as given by Theorems 3.1.B, 3.1.C, and 3.1.E. Numerical techniques are explored in Part III of Gentle's book.

Note. On page 58, Gentle gives an argument that the area of the parallelogram determined by two 2 -dimensional vectors can be calculated by taking the determinant of the $2 \times 2$ matrix with the vectors as columns. In Section 4.1, "Areas, Volumes, and Cross Products," of Fraleigh and Beauregard's Linear Algebra, 3rd edition (Addison-Wesley, 1995), it is shown that the volume of the box determined by three 3 -dimensional vectors can be calculated by taking the determinant of the $3 \times 3$ matrix with the vectors as columns. For details, see my online notes on 4.2. Areas, Volumes, and Cross Products.

Note. Given the way determinants are presented in both sophomore linear algebra and in this class, it is surprising to realize that the concept of determinant predates that of matrix! This is because determinants were introduced in order to solve systems of equations by Gottfried Wilhelm Leibniz (1646-1716) in 1693, though his approach remained unknown at the time [Issrael Kleiner, A History of Abstract Algebra, Birkhäuser (2007), page 80]. You might be familiar with "Cramer's Rule" which gives the solution to a system of $n$ equations in $n$ unknowns, $A \vec{x}=\vec{b}$, where $A^{-1}$ exists as $x_{k}=\operatorname{det}\left(B_{k}\right) / \operatorname{det}(A)$ for $k=1,2, \ldots, n$ where $B_{k}$ is the matrix
obtained from $A$ by replacing the $k$ th column vector of $A$ by the vector $\vec{b}$ (see Fraleigh and Beauregard's Linear Algebra, 3rd Edition, Section 4.3, "Computation of Determinants and Cramer's Rule"; see also Theorem 4.5 on page 2 of these online notes). Gabriel Cramer (1704-1752) published this result in his Introduction to the Analysis of Algebraic Curves in 1750, but gave no proof (Fraleigh and Beauregard and my online class notes include a proof). The first publication to contain some information on determinant was Colin Maclaurin's (1698-1746) 1748 Treatise in Algebra in which he used determinants to solve $2 \times 2$ and $3 \times 3$ systems [Kleiner, page 81]. Augustin Cauchy (1789-1857) gave the first systematic study of determinants in 1815 in his paper "On the Functions which Can Assume But Two Equal Values of Opposite Sign by Means of Transformations Carried Out to Their Variables," Journal de l'Ecole Polytechique 10, 29-112. In this early work, he proves $\operatorname{det}(A B)=$ $\operatorname{det}(A) \operatorname{det}(B)$.


Karl Weierstrass (1815-1897) and Leopold Kronecker (1823-1891) gave a definition of determinant in terms of axioms, probably in the 1860s [Kleiner, page 81]. By 1880, most of the basic results of linear algebra had been established, but not in a cohesive theory. In 1888, Giuseppe Peano (1858-1932) introduced the idea of a vector space, taking a large step forward in the development of a general theory,
which was to follow in the early decades of the twentieth century [Kleiner, page 79]. Images are from the MacTutor History of Mathematics archive.

Revised: 6/15/2020

