CHAPTER 6

Determinants

6.1 DETERMINANTS

At the beginning of this text, reference was made to the ancient Chinese counting board on which colored bamboo rods were manipulated according to prescribed "rules of thumb" in order to solve a system of linear equations. The Chinese counting board is believed to date back to at least 200 B.C., and it was used more or less in the same way for a millennium. The counting board and the "rules of thumb" eventually found their way to Japan where Seki Kowa (1642–1708), a great Japanese mathematician, synthesized the ancient Chinese ideas of array manipulation. Kowa formulated the concept of what we now call the determinant to facilitate solving linear systems—his definition is thought to have been made some time before 1683.

About the same time—somewhere between 1678 and 1693—Gottfried W. Leibniz (1646–1716), a German mathematician, was independently developing his own concept of the determinant together with applications of array manipulation to solve systems of linear equations. It appears that Leibniz's early work dealt with only three equations in three unknowns, whereas Seki Kowa gave a general treatment for n equations in n unknowns. It seems that Kowa and Leibniz both developed what later became known as Cramer's rule (p. 476), but not in the same form or notation. These men had something else in common—their ideas concerning the solution of linear systems were never adopted by the mathematical community of their time, and their discoveries quickly faded into oblivion.

Eventually the determinant was rediscovered, and much was written on the subject between 1750 and 1900. During this era, determinants became the major tool used to analyze and solve linear systems, while the theory of matrices remained relatively undeveloped. But mathematics, like a river, is everchanging in its course, and major branches can dry up to become minor tributaries while small trickling brooks can develop into raging torrents. This is precisely what occurred with determinants and matrices. The study and use of determinants eventually gave way to Cayley's matrix algebra, and today matrix and linear algebra are in the main stream of applied mathematics, while the role of determinants has been relegated to a minor backwater position. Nevertheless, it is still important to understand what a determinant is and to learn a few of its fundamental properties. Our goal is not to study determinants for their own sake, but rather to explore those properties that are useful in the further development of matrix theory and its applications. Accordingly, many secondary properties are omitted or confined to the exercises, and the details in proofs will be kept to a minimum.

Over the years there have evolved various "slick" ways to define the determinant, but each of these "slick" approaches seems to require at least one "sticky" theorem in order to make the theory sound. We are going to opt for expedience over elegance and proceed with the classical treatment.

A ***permutation*** $p = (p_1, p_2, \ldots, p_n)$ of the numbers $(1, 2, \ldots, n)$ is simply any rearrangement. For example, the set

$$\{(1,2,3) \quad (1,3,2) \quad (2,1,3) \quad (2,3,1) \quad (3,1,2) \quad (3,2,1)\}$$

contains the six distinct permutations of $(1,2,3)$. In general, the sequence $(1,2,\ldots,n)$ has $n! = n(n-1)(n-2)\cdots 1$ different permutations. Given a permutation, consider the problem of restoring it to natural order by a sequence of pairwise interchanges. For example, $(1,4,3,2)$ can be restored to natural order with a single interchange of 2 and 4 or, as indicated in Figure 6.1.1, three *adjacent* interchanges can be used.

FIGURE 6.1.1

The important thing here is that both 1 and 3 are odd. Try to restore $(1,4,3,2)$ to natural order by using an even number of interchanges, and you will discover that it is impossible. This is due to the following general rule that is stated without proof. *The* ***parity*** *of a permutation is unique*—i.e., if a permutation p can be restored to natural order by an even (odd) number of interchanges, then every other sequence of interchanges that restores p to natural order must also be even (odd). Accordingly, the ***sign of a permutation*** p is defined to be the number

$$\sigma(p) = \begin{cases} +1 & \text{if } p \text{ can be restored to natural order by an} \\ & \textit{even} \text{ number of interchanges,} \\ -1 & \text{if } p \text{ can be restored to natural order by an} \\ & \textit{odd} \text{ number of interchanges.} \end{cases}$$

For example, if $p = (1,4,3,2)$, then $\sigma(p) = -1$, and if $p = (4,3,2,1)$, then $\sigma(p) = +1$. The sign of the natural order $p = (1,2,3,4)$ is naturally $\sigma(p) = +1$. The general definition of the determinant can now be given.

Definition of Determinant

For an $n \times n$ matrix $\mathbf{A} = [a_{ij}]$, the ***determinant*** of $\mathbf{A}$ is defined to be the scalar

$$\det(\mathbf{A}) = \sum_{p} \sigma(p) a_{1p_1} a_{2p_2} \cdots a_{np_n}, \tag{6.1.1}$$

where the sum is taken over the $n!$ permutations $p = (p_1, p_2, \ldots, p_n)$ of $(1, 2, \ldots, n)$. Observe that each term $a_{1p_1} a_{2p_2} \cdots a_{np_n}$ in (6.1.1) contains exactly one entry from each row and each column of $\mathbf{A}$. The determinant of $\mathbf{A}$ can be denoted by $\det(\mathbf{A})$ or $|\mathbf{A}|$, whichever is more convenient.

Note: The determinant of a nonsquare matrix is not defined.

For example, when $\mathbf{A}$ is 2×2 there are $2! = 2$ permutations of (1,2), namely, $\{(1,2) \quad (2,1)\}$, so $\det(\mathbf{A})$ contains the two terms

$$\sigma(1,2) a_{11} a_{22} \quad \text{and} \quad \sigma(2,1) a_{12} a_{21}.$$

Since $\sigma(1,2) = +1$ and $\sigma(2,1) = -1$, we obtain the familiar formula

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} a_{22} - a_{12} a_{21}. \tag{6.1.2}$$

Example 6.1.1

Problem: Use the definition to compute $\det(\mathbf{A})$, where $\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$.

Solution: The $3! = 6$ permutations of $(1, 2, 3)$ together with the terms in the expansion of $\det(\mathbf{A})$ are shown in Table 6.1.1.

TABLE 6.1.1

$p = (p_1, p_2, p_3)$	$\sigma(p)$	$a_{1p_1}a_{2p_2}a_{3p_3}$
$(1, 2, 3)$	$+$	$1 \times 5 \times 9 = 45$
$(1, 3, 2)$	$-$	$1 \times 6 \times 8 = 48$
$(2, 1, 3)$	$-$	$2 \times 4 \times 9 = 72$
$(2, 3, 1)$	$+$	$2 \times 6 \times 7 = 84$
$(3, 1, 2)$	$+$	$3 \times 4 \times 8 = 96$
$(3, 2, 1)$	$-$	$3 \times 5 \times 7 = 105$

Therefore,

$$\det(\mathbf{A}) = \sum_p \sigma(p)a_{1p_1}a_{2p_2}a_{3p_3} = 45 - 48 - 72 + 84 + 96 - 105 = 0.$$

Perhaps you have seen rules for computing 3×3 determinants that involve running up, down, and around various diagonal lines. These rules do not easily generalize to matrices of order greater than three, and in case you have forgotten (or never knew) them, do not worry about it. Remember the 2×2 rule given in (6.1.2) as well as the following statement concerning triangular matrices and let it go at that.

Triangular Determinants

The determinant of a triangular matrix is the product of its diagonal entries. In other words,

$$\begin{vmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t_{nn} \end{vmatrix} = t_{11}t_{22}\cdots t_{nn}. \tag{6.1.3}$$

Proof. Recall from the definition (6.1.1) that each term $t_{1p_1}t_{2p_2}\cdots t_{np_n}$ contains exactly one entry from each row and each column. This means that there is only one term in the expansion of the determinant that does not contain an entry below the diagonal, and this term is $t_{11}t_{22}\cdots t_{nn}$. ∎

Transposition Doesn't Alter Determinants

- $\det\left(\mathbf{A}^T\right) = \det(\mathbf{A})$ for all $n \times n$ matrices. (6.1.4)

Proof. As $p = (p_1, p_2, \ldots, p_n)$ varies over all permutations of $(1, 2, \ldots, n)$, the set of all products $\{\sigma(p) a_{1p_1} a_{2p_2} \cdots a_{np_n}\}$ is the same as the set of all products $\{\sigma(p) a_{p_1 1} a_{p_2 2} \cdots a_{p_n n}\}$. Explicitly construct both of these sets for $n = 3$ to convince yourself. ■

Equation (6.1.4) insures that it's not necessary to distinguish between rows and columns when discussing properties of determinants, so theorems concerning determinants that involve row manipulations will remain true when the word "row" is replaced by "column." For example, it's essential to know how elementary row and column operations alter the determinant of a matrix, but, by virtue of (6.1.4), it suffices to limit the discussion to elementary row operations.

Effects of Row Operations

Let $\mathbf{B}$ be the matrix obtained from $\mathbf{A}_{n\times n}$ by one of the three elementary row operations:

Type I: Interchange rows i and j.

Type II: Multiply row i by $\alpha \neq 0$.

Type III: Add α times row i to row j.

The value of $\det(\mathbf{B})$ is as follows:

- $\det(\mathbf{B}) = -\det(\mathbf{A})$ for Type I operations. (6.1.5)
- $\det(\mathbf{B}) = \alpha \det(\mathbf{A})$ for Type II operations. (6.1.6)
- $\det(\mathbf{B}) = \det(\mathbf{A})$ for Type III operations. (6.1.7)

Proof of (6.1.5). If $\mathbf{B}$ agrees with $\mathbf{A}$ except that $\mathbf{B}_{i*} = \mathbf{A}_{j*}$ and $\mathbf{B}_{j*} = \mathbf{A}_{i*}$, then for each permutation $p = (p_1, p_2, \ldots, p_n)$ of $(1, 2, \ldots, n)$,

$$
\begin{aligned}
b_{1p_1} \cdots b_{ip_i} \cdots b_{jp_j} \cdots b_{np_n} &= a_{1p_1} \cdots a_{jp_i} \cdots a_{ip_j} \cdots a_{np_n} \\
&= a_{1p_1} \cdots a_{ip_j} \cdots a_{jp_i} \cdots a_{np_n}.
\end{aligned}
$$

Furthermore, $\sigma(p_1, \ldots, p_i, \ldots, p_j, \ldots, p_n) = -\sigma(p_1, \ldots, p_j, \ldots, p_i, \ldots, p_n)$ because the two permutations differ only by one interchange. Consequently, definition (6.1.1) of the determinant guarantees that $\det(\mathbf{B}) = -\det(\mathbf{A})$.

Proof of (6.1.6). If $\mathbf{B}$ agrees with $\mathbf{A}$ except that $\mathbf{B}_{i*} = \alpha\mathbf{A}_{i*}$, then for each permutation $p = (p_1, p_2, \ldots, p_n)$,

$$b_{1p_1} \cdots b_{ip_i} \cdots b_{np_n} = a_{1p_1} \cdots \alpha a_{ip_i} \cdots a_{np_n} = \alpha(a_{1p_1} \cdots a_{ip_i} \cdots a_{np_n}),$$

and therefore the expansion (6.1.1) yields $\det(\mathbf{B}) = \alpha \det(\mathbf{A})$.

Proof of (6.1.7). If $\mathbf{B}$ agrees with $\mathbf{A}$ except that $\mathbf{B}_{j*} = \mathbf{A}_{j*} + \alpha\mathbf{A}_{i*}$, then for each permutation $p = (p_1, p_2, \ldots, p_n)$,

$$\begin{aligned} b_{1p_1} \cdots b_{ip_i} \cdots b_{jp_j} \cdots b_{np_n} &= a_{1p_1} \cdots a_{ip_i} \cdots (a_{jp_j} + \alpha a_{ip_j}) \cdots a_{np_n} \\ &= a_{1p_1} \cdots a_{ip_i} \cdots a_{jp_j} \cdots a_{np_n} + \alpha(a_{1p_1} \cdots a_{ip_i} \cdots a_{ip_j} \cdots a_{np_n}), \end{aligned}$$

so that

$$\begin{aligned} \det(\mathbf{B}) = &\sum_p \sigma(p) a_{1p_1} \cdots a_{ip_i} \cdots a_{jp_j} \cdots a_{np_n} \\ &+ \alpha \sum_p \sigma(p) a_{1p_1} \cdots a_{ip_i} \cdots a_{ip_j} \cdots a_{np_n}. \end{aligned} \tag{6.1.8}$$

The first sum on the right-hand side of (6.1.8) is $\det(\mathbf{A})$, while the second sum is the expansion of the determinant of a matrix $\tilde{\mathbf{A}}$ in which the i^{th} and j^{th} rows are identical. For such a matrix, $\det(\tilde{\mathbf{A}}) = 0$ because (6.1.5) says that the sign of the determinant is reversed whenever the i^{th} and j^{th} rows are interchanged, so $\det(\tilde{\mathbf{A}}) = -\det(\tilde{\mathbf{A}})$. Consequently, the second sum on the right-hand side of (6.1.8) is zero, and thus $\det(\mathbf{B}) = \det(\mathbf{A})$. ∎

It is now possible to evaluate the determinant of an elementary matrix associated with any of the three types of elementary operations. Let $\mathbf{E}$, $\mathbf{F}$, and $\mathbf{G}$ be elementary matrices of Types I, II, and III, respectively, and recall from the discussion in §3.9 that each of these elementary matrices can be obtained by performing the associated row (or column) operation to an identity matrix of appropriate size. The result concerning triangular determinants (6.1.3) guarantees that $\det(\mathbf{I}) = 1$ regardless of the size of $\mathbf{I}$, so if $\mathbf{E}$ is obtained by interchanging any two rows (or columns) in $\mathbf{I}$, then (6.1.5) insures that

$$\det(\mathbf{E}) = -\det(\mathbf{I}) = -1. \tag{6.1.9}$$

Similarly, if $\mathbf{F}$ is obtained by multiplying any row (or column) in $\mathbf{I}$ by $\alpha \neq 0$, then (6.1.6) implies that

$$\det(\mathbf{F}) = \alpha \det(\mathbf{I}) = \alpha, \tag{6.1.10}$$

and if $\mathbf{G}$ is the result of adding a multiple of one row (or column) in $\mathbf{I}$ to another row (or column) in $\mathbf{I}$, then (6.1.7) guarantees that

$$\det(\mathbf{G}) = \det(\mathbf{I}) = 1. \tag{6.1.11}$$

In particular, (6.1.9)–(6.1.11) guarantee that the determinants of elementary matrices of Types I, II, and III are nonzero.

As discussed in §3.9, if $\mathbf{P}$ is an elementary matrix of Type I, II, or III, and if $\mathbf{A}$ is any other matrix, then the product $\mathbf{PA}$ is the matrix obtained by performing the elementary operation associated with $\mathbf{P}$ to the rows of $\mathbf{A}$. This, together with the observations (6.1.5)–(6.1.7) and (6.1.9)–(6.1.11), leads to the conclusion that for every square matrix $\mathbf{A}$,

$$\begin{aligned}
\det(\mathbf{EA}) &= -\det(\mathbf{A}) &&= \det(\mathbf{E})\det(\mathbf{A}),\\
\det(\mathbf{FA}) &= \alpha\det(\mathbf{A}) &&= \det(\mathbf{F})\det(\mathbf{A}),\\
\det(\mathbf{GA}) &= \det(\mathbf{A}) &&= \det(\mathbf{G})\det(\mathbf{A}).
\end{aligned}$$

In other words, $\det(\mathbf{PA}) = \det(\mathbf{P})\det(\mathbf{A})$ whenever $\mathbf{P}$ is an elementary matrix of Type I, II, or III. It's easy to extend this observation to any number of these elementary matrices, $\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_k$, by writing

$$\begin{aligned}
\det(\mathbf{P}_1\mathbf{P}_2\cdots\mathbf{P}_k\mathbf{A}) &= \det(\mathbf{P}_1)\det(\mathbf{P}_2\cdots\mathbf{P}_k\mathbf{A})\\
&= \det(\mathbf{P}_1)\det(\mathbf{P}_2)\det(\mathbf{P}_3\cdots\mathbf{P}_k\mathbf{A})\\
&\;\;\vdots\\
&= \det(\mathbf{P}_1)\det(\mathbf{P}_2)\cdots\det(\mathbf{P}_k)\det(\mathbf{A}).
\end{aligned} \tag{6.1.12}$$

This leads to a characterization of invertibility in terms of determinants.

Invertibility and Determinants

- $\mathbf{A}_{n\times n}$ is nonsingular if and only if $\det(\mathbf{A}) \neq 0$ (6.1.13)

  or, equivalently,

- $\mathbf{A}_{n\times n}$ is singular if and only if $\det(\mathbf{A}) = 0$. (6.1.14)

Proof. Let $\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_k$ be a sequence of elementary matrices of Type I, II, or III such that $\mathbf{P}_1\mathbf{P}_2\cdots\mathbf{P}_k\mathbf{A} = \mathbf{E}_\mathbf{A}$, and apply (6.1.12) to conclude

$$\det(\mathbf{P}_1)\det(\mathbf{P}_2)\cdots\det(\mathbf{P}_k)\det(\mathbf{A}) = \det(\mathbf{E}_\mathbf{A}).$$

Since elementary matrices have nonzero determinants,

$$\begin{aligned}
\det(\mathbf{A}) \neq 0 \iff \det(\mathbf{E}_\mathbf{A}) \neq 0 &\iff \text{there are no zero pivots}\\
&\iff \text{every column in } \mathbf{E}_\mathbf{A} \text{ (and in } \mathbf{A}\text{) is basic}\\
&\iff \mathbf{A} \text{ is nonsingular.} \quad \blacksquare
\end{aligned}$$

Example 6.1.2

Caution! Small Determinants $\not\Longleftrightarrow$ Near Singularity. Because of (6.1.13) and (6.1.14), it might be easy to get the idea that $\det(\mathbf{A})$ is somehow a measure of how close $\mathbf{A}$ is to being singular, but this is not necessarily the case. Nearly singular matrices need not have determinants of small magnitude. For example, $\mathbf{A}_n = \begin{pmatrix} n & 0 \\ 0 & 1/n \end{pmatrix}$ is nearly singular when n is large, but $\det(\mathbf{A}_n) = 1$ for all n. Furthermore, small determinants do not necessarily signal nearly singular matrices. For example,

$$\mathbf{A}_n = \begin{pmatrix} .1 & 0 & \cdots & 0 \\ 0 & .1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & .1 \end{pmatrix}_{n\times n}$$

is not close to any singular matrix—see (5.12.10) on p. 417—but $\det(\mathbf{A}_n) = (.1)^n$ is extremely small for large n.

A ***minor determinant*** (or simply a ***minor***) of $\mathbf{A}_{m\times n}$ is defined to be the determinant of any $k \times k$ submatrix of $\mathbf{A}$. For example,

$$\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} = -3 \text{ and } \begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} = -6 \text{ are } 2\times 2 \text{ minors of } \mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

An individual entry of $\mathbf{A}$ can be regarded as a 1×1 minor, and $\det(\mathbf{A})$ itself is considered to be a 3×3 minor of $\mathbf{A}$.

We already know that the rank of any matrix $\mathbf{A}$ is the size of the largest nonsingular submatrix in $\mathbf{A}$ (p. 215). But (6.1.13) guarantees that the nonsingular submatrices of $\mathbf{A}$ are simply those submatrices with nonzero determinants, so we have the following characterization of rank.

Rank and Determinants

- $rank(\mathbf{A})$ = the size of the largest nonzero minor of $\mathbf{A}$.

Example 6.1.3

Problem: Use determinants to compute the rank of $\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 1 \\ 7 & 8 & 9 & 1 \end{pmatrix}$.

Solution: Clearly, there are 1×1 and 2×2 minors that are nonzero, so $rank(\mathbf{A}) \geq 2$. In order to decide if the rank is three, we must see if there are any 3×3 nonzero minors. There are exactly four 3×3 minors, and they are

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 2 & 1 \\ 4 & 5 & 1 \\ 7 & 8 & 1 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 3 & 1 \\ 4 & 6 & 1 \\ 7 & 9 & 1 \end{vmatrix} = 0, \quad \begin{vmatrix} 2 & 3 & 1 \\ 5 & 6 & 1 \\ 8 & 9 & 1 \end{vmatrix} = 0.$$

Since all 3×3 minors are 0, we conclude that $rank\,(\mathbf{A}) = 2$. You should be able to see from this example that using determinants is generally not a good way to compute the rank of a matrix.

In (6.1.12) we observed that the determinant of a product of elementary matrices is the product of their respective determinants. We are now in a position to extend this observation.

Product Rules

- $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$ for all $n \times n$ matrices. (6.1.15)
- $\det\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} \end{pmatrix} = \det(\mathbf{A})\det(\mathbf{D})$ if $\mathbf{A}$ and $\mathbf{D}$ are square. (6.1.16)

Proof of (6.1.15). If $\mathbf{A}$ is singular, then $\mathbf{AB}$ is also singular because (4.5.2) says that $rank\,(\mathbf{AB}) \le rank\,(\mathbf{A})$. Consequently, (6.1.14) implies that

$$\det(\mathbf{AB}) = 0 = \det(\mathbf{A})\det(\mathbf{B}),$$

so (6.1.15) is trivially true when $\mathbf{A}$ is singular. If $\mathbf{A}$ is nonsingular, then $\mathbf{A}$ can be written as a product of elementary matrices $\mathbf{A} = \mathbf{P}_1\mathbf{P}_2\cdots\mathbf{P}_k$ that are of Type I, II, or III—recall (3.9.3). Therefore, (6.1.12) can be applied to produce

$$\begin{aligned} \det(\mathbf{AB}) &= \det(\mathbf{P}_1\mathbf{P}_2\cdots\mathbf{P}_k\mathbf{B}) = \det(\mathbf{P}_1)\det(\mathbf{P}_2)\cdots\det(\mathbf{P}_k)\det(\mathbf{B}) \\ &= \det(\mathbf{P}_1\mathbf{P}_2\cdots\mathbf{P}_k)\det(\mathbf{B}) = \det(\mathbf{A})\det(\mathbf{B}). \end{aligned}$$

Proof of (6.1.16). First consider the special case $\mathbf{X} = \begin{pmatrix} \mathbf{A}_{r\times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$, and use the definition to write $\det(\mathbf{X}) = \sum_{\sigma(p)} x_{1j_1}x_{2j_2}\cdots x_{rj_r}x_{r+1,j_{r+1}}\cdots x_{n,j_n}$. But

$$x_{rj_r}x_{r+1,j_{r+1}}\cdots x_{n,j_n} = \begin{cases} 1 & \text{when } p = \begin{pmatrix} 1 & \cdots & r & r+1 & \cdots & n \\ j_1 & \cdots & j_r & r+1 & \cdots & n \end{pmatrix}, \\ 0 & \text{for all other permutations,} \end{cases}$$

so, if p_r denotes permutations of only the first r positive integers, then

$$\det(\mathbf{X}) = \sum_{\sigma(p)} x_{1j_1}x_{2j_2}\cdots x_{rj_r}x_{r+1,j_{r+1}}\cdots x_{n,j_n} = \sum_{\sigma(p_r)} x_{1j_1}x_{2j_2}\cdots x_{rj_r} = \det(\mathbf{A}).$$

Thus $\begin{vmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{vmatrix} = \det(\mathbf{A})$. Similarly, $\begin{vmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{vmatrix} = \det(\mathbf{D})$, so, by (6.1.15),

$$\begin{vmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{vmatrix} = \det\left\{\begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{pmatrix}\right\} = \begin{vmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{vmatrix}\begin{vmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{vmatrix} = \det(\mathbf{A})\det(\mathbf{D}).$$

If $\mathbf{A} = \mathbf{Q}_A\mathbf{R}_A$ and $\mathbf{D} = \mathbf{Q}_D\mathbf{R}_D$ are the respective QR factorizations (p. 345) of $\mathbf{A}$ and $\mathbf{D}$, then $\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{Q}_A & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_D \end{pmatrix}\begin{pmatrix} \mathbf{R}_A & \mathbf{Q}_A^T\mathbf{B} \\ \mathbf{0} & \mathbf{R}_D \end{pmatrix}$ is also a QR factorization. By (6.1.3), the determinant of a triangular matrix is the product of its diagonal entries, and this together with the previous results yield

$$\begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} \end{vmatrix} = \begin{vmatrix} \mathbf{Q}_A & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_D \end{vmatrix}\begin{vmatrix} \mathbf{R}_A & \mathbf{Q}_A^T\mathbf{B} \\ \mathbf{0} & \mathbf{R}_D \end{vmatrix} = \det(\mathbf{Q}_A)\det(\mathbf{Q}_D)\det(\mathbf{R}_A)\det(\mathbf{R}_D)$$
$$= \det(\mathbf{Q}_A\mathbf{R}_A)\det(\mathbf{Q}_D\mathbf{R}_D) = \det(\mathbf{A})\det(\mathbf{D}). \quad \blacksquare$$

Example 6.1.4

Volume and Determinants. The definition of a determinant is purely algebraic, but there is a concrete geometrical interpretation. A solid in $\Re^m$ with parallel opposing faces whose adjacent sides are defined by vectors from a linearly independent set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is called an n-dimensional ***parallelepiped***. As depicted in Figure 6.1.2, a two-dimensional parallelepiped is a parallelogram, and a three-dimensional parallelepiped is a skewed rectangular box.

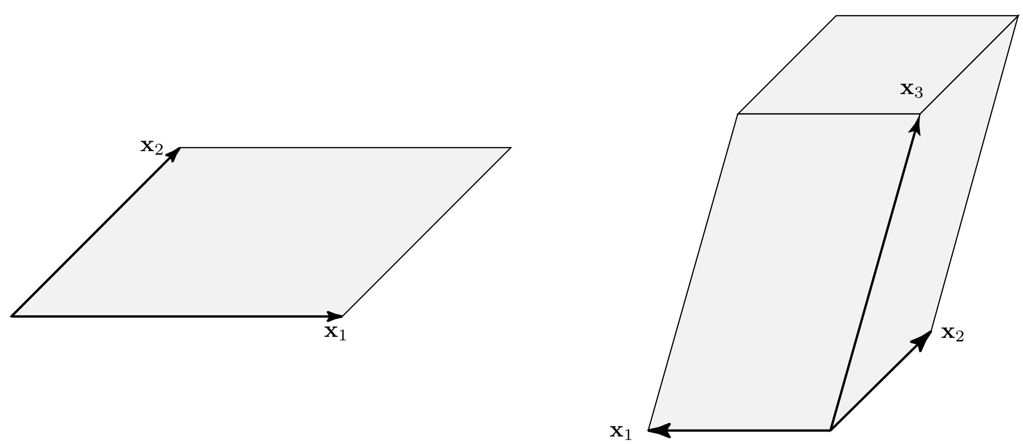

FIGURE 6.1.2

Problem: When $\mathbf{A} \in \Re^{m\times n}$ has linearly independent columns, explain why the volume of the n-dimensional parallelepiped generated by the columns of $\mathbf{A}$ is $V_n = \left[\det\left(\mathbf{A}^T\mathbf{A}\right)\right]^{1/2}$. In particular, if $\mathbf{A}$ is square, then $V_n = |\det(\mathbf{A})|$.

Solution: Recall from Example 5.13.2 on p. 431 that if $\mathbf{A}_{m\times n} = \mathbf{Q}_{m\times n}\mathbf{R}_{n\times n}$ is the (rectangular) QR factorization of $\mathbf{A}$, then the volume of the n-dimensional parallelepiped generated by the columns of $\mathbf{A}$ is $V_n = \nu_1\nu_2\cdots\nu_n = \det(\mathbf{R})$, where the ν_k 's are the diagonal elements of the upper-triangular matrix $\mathbf{R}$. Use

$\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$ together with the product rule (6.1.15) and the fact that transposition doesn't affect determinants (6.1.4) to write

$$\begin{aligned}\det\left(\mathbf{A}^T\mathbf{A}\right) &= \det\left(\mathbf{R}^T\mathbf{Q}^T\mathbf{Q}\mathbf{R}\right) = \det\left(\mathbf{R}^T\mathbf{R}\right) = \det\left(\mathbf{R}^T\right)\det\left(\mathbf{R}\right)\\ &= (\det\left(\mathbf{R}\right))^2 = (\nu_1\nu_2\cdots\nu_n)^2 = V_n^2.\end{aligned} \tag{6.1.17}$$

If $\mathbf{A}$ is square, $\det\left(\mathbf{A}^T\mathbf{A}\right) = \det\left(\mathbf{A}^T\right)\det\left(\mathbf{A}\right) = (\det\left(\mathbf{A}\right))^2$, so $V_n = |\det\left(\mathbf{A}\right)|$.

Hadamard's Inequality: Recall from (5.13.7) that if

$$\mathbf{A} = \left[\mathbf{x}_1 \,|\, \mathbf{x}_2 \,|\, \cdots \,|\, \mathbf{x}_n\right]_{n\times n} \quad \text{and} \quad \mathbf{A}_j = \left[\mathbf{x}_1 \,|\, \mathbf{x}_2 \,|\, \cdots \,|\, \mathbf{x}_j\right]_{n\times j},$$

then $\nu_1 = \|\mathbf{x}_1\|_2$ and $\nu_k = \|(\mathbf{I} - \mathbf{P}_k)\mathbf{x}_k\|_2$ (the projected height of $\mathbf{x}_k$) for $k > 1$, where $\mathbf{P}_k$ is the orthogonal projector onto $R\left(\mathbf{A}_{k-1}\right)$. But

$$\nu_k^2 = \|(\mathbf{I} - \mathbf{P}_k)\mathbf{x}_k\|_2^2 \le \|(\mathbf{I} - \mathbf{P}_k)\|_2^2\,\|\mathbf{x}_k\|_2^2 = \|\mathbf{x}_k\|_2^2 \quad \text{(recall (5.13.10))},$$

so, by (6.1.17), $\det\left(\mathbf{A}^T\mathbf{A}\right) \le \|\mathbf{x}_1\|_2^2\,\|\mathbf{x}_2\|_2^2\cdots\|\mathbf{x}_n\|_2^2$ or, equivalently,

$$|\det\left(\mathbf{A}\right)| \le \prod_{k=1}^{n}\|\mathbf{x}_k\|_2 = \prod_{j=1}^{n}\left(\sum_{i=1}^{n}|a_{ij}|^2\right)^{1/2}, \tag{6.1.18}$$

with equality holding if and only if the $\mathbf{x}_k$'s are mutually orthogonal. This is ***Hadamard's inequality.***[64] In light of the preceding discussion, it simply asserts that the volume of the parallelepiped $\mathcal{P}$ generated by the columns of $\mathbf{A}$ can't exceed the volume of a rectangular box whose sides have length $\|\mathbf{x}_k\|_2$, a fact that is geometrically evident because $\mathcal{P}$ is a *skewed* rectangular box with sides of length $\|\mathbf{x}_k\|_2$.

The product rule (6.1.15) provides a practical way to compute determinants. Recall from §3.10 that for every nonsingular matrix $\mathbf{A}$, there is a permutation matrix $\mathbf{P}$ (which is a product of elementary interchange matrices) such that $\mathbf{PA} = \mathbf{LU}$ in which $\mathbf{L}$ is lower triangular with 1's on its diagonal, and $\mathbf{U}$ is upper triangular with the pivots on its diagonal. The product rule guarantees

[64] Jacques Hadamard (1865–1963), a leading French mathematician of the first half of the twentieth century, discovered this inequality in 1893. Influenced in part by the tragic death of his sons in World War I, Hadamard became a peace activist whose politics drifted far left to the extent that the United States was reluctant to allow him to enter the country to attend the International Congress of Mathematicians held in Cambridge, Massachusetts, in 1950. Due to support from influential mathematicians, Hadamard was made honorary president of the congress, and the resulting visibility together with pressure from important U.S. scientists forced officials to allow him to attend.

that $\det(\mathbf{P})\det(\mathbf{A}) = \det(\mathbf{L})\det(\mathbf{U})$, and we know from (6.1.9) that if $\mathbf{E}$ is an elementary interchange matrix, then $\det(\mathbf{E}) = -1$, so

$$\det(\mathbf{P}) = \begin{cases} +1 & \text{if } \mathbf{P} \text{ is the product of an } \textit{even} \text{ number of interchanges,} \\ -1 & \text{if } \mathbf{P} \text{ is the product of an } \textit{odd} \text{ number of interchanges.} \end{cases}$$

The result concerning triangular determinants (6.1.3) shows that $\det(\mathbf{L}) = 1$ and $\det(\mathbf{U}) = u_{11}u_{22}\cdots u_{nn}$, where the u_{ii}'s are the pivots, so, putting these observations together yields $\det(\mathbf{A}) = \pm u_{11}u_{22}\cdots u_{nn}$, where the sign depends on the number of row interchanges used. Below is a summary.

Computing a Determinant

If $\mathbf{PA}_{n\times n} = \mathbf{LU}$ is an LU factorization obtained with row interchanges (use partial pivoting for numerical stability), then

$$\det(\mathbf{A}) = \sigma u_{11}u_{22}\cdots u_{nn}.$$

The u_{ii}'s are the pivots, and σ is the sign of the permutation. That is,

$$\sigma = \begin{cases} +1 & \text{if an } \textit{even} \text{ number of row interchanges are used,} \\ -1 & \text{if an } \textit{odd} \text{ number of row interchanges are used.} \end{cases}$$

If a zero pivot emerges that cannot be removed (because all entries below the pivot are zero), then $\mathbf{A}$ is singular and $\det(\mathbf{A}) = 0$. Exercise 6.2.18 discusses orthogonal reduction to compute $\det(\mathbf{A})$.

Example 6.1.5

Problem: Use partial pivoting to determine an LU decomposition $\mathbf{PA} = \mathbf{LU}$, and then evaluate the determinant of $\mathbf{A} = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 4 & 8 & 12 & -8 \\ 2 & 3 & 2 & 1 \\ -3 & -1 & 1 & -4 \end{pmatrix}$.

Solution: The LU factors of $\mathbf{A}$ were computed in Example 3.10.4 as follows.

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -3/4 & 1 & 0 & 0 \\ 1/4 & 0 & 1 & 0 \\ 1/2 & -1/5 & 1/3 & 1 \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} 4 & 8 & 12 & -8 \\ 0 & 5 & 10 & -10 \\ 0 & 0 & -6 & 6 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{P} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

The only modification needed is to keep track of how many row interchanges are used. Reviewing Example 3.10.4 reveals that the pivoting process required three interchanges, so $\sigma = -1$, and hence $\det(\mathbf{A}) = (-1)(4)(5)(-6)(1) = 120$.

It's sometimes necessary to compute the derivative of a determinant whose entries are differentiable functions. The following formula shows how this is done.

Derivative of a Determinant

If the entries in $\mathbf{A}_{n\times n} = [a_{ij}(t)]$ are differentiable functions of t, then

$$\frac{d\big(\det(\mathbf{A})\big)}{dt} = \det(\mathbf{D}_1) + \det(\mathbf{D}_2) + \cdots + \det(\mathbf{D}_n), \tag{6.1.19}$$

where $\mathbf{D}_i$ is identical to $\mathbf{A}$ except that the entries in the i^{th} row are replaced by their derivatives—i.e., $[\mathbf{D}_i]_{k*} = \begin{cases} \mathbf{A}_{k*} & \text{if } i \neq k, \\ d\,\mathbf{A}_{k*}/dt & \text{if } i = k. \end{cases}$

Proof. This follows directly from the definition of a determinant by writing

$$\begin{aligned}
\frac{d\big(\det(\mathbf{A})\big)}{dt} &= \frac{d}{dt}\sum_p \sigma(p) a_{1p_1}a_{2p_2}\cdots a_{np_n} = \sum_p \sigma(p)\frac{d\big(a_{1p_1}a_{2p_2}\cdots a_{np_n}\big)}{dt} \\
&= \sum_p \sigma(p)\Big(a'_{1p_1}a_{2p_2}\cdots a_{np_n} + a_{1p_1}a'_{2p_2}\cdots a_{np_n} + \cdots + a_{1p_1}a_{2p_2}\cdots a'_{np_n}\Big) \\
&= \sum_p \sigma(p)a'_{1p_1}a_{2p_2}\cdots a_{np_n} + \sum_p \sigma(p)a_{1p_1}a'_{2p_2}\cdots a_{np_n} \\
&\qquad + \cdots + \sum_p \sigma(p)a_{1p_1}a_{2p_2}\cdots a'_{np_n} \\
&= \det(\mathbf{D}_1) + \det(\mathbf{D}_2) + \cdots + \det(\mathbf{D}_n). \quad \blacksquare
\end{aligned}$$

Example 6.1.6

Problem: Evaluate the derivative $d\big(\det(\mathbf{A})\big)/dt$ for $\mathbf{A} = \begin{pmatrix} e^t & e^{-t} \\ \cos t & \sin t \end{pmatrix}$.

Solution: Applying formula (6.1.19) yields

$$\frac{d\big(\det(\mathbf{A})\big)}{dt} = \begin{vmatrix} e^t & -e^{-t} \\ \cos t & \sin t \end{vmatrix} + \begin{vmatrix} e^t & e^{-t} \\ -\sin t & \cos t \end{vmatrix} = \left(e^t + e^{-t}\right)(\cos t + \sin t).$$

Check this by first expanding $\det(\mathbf{A})$ and then computing the derivative.

Exercises for section 6.1

6.1.1. Use the definition to evaluate $\det(\mathbf{A})$ for each of the following matrices.

(a) $\mathbf{A} = \begin{pmatrix} 3 & -2 & 1 \\ -5 & 4 & 0 \\ 2 & 1 & 6 \end{pmatrix}$. (b) $\mathbf{A} = \begin{pmatrix} 2 & 1 & 1 \\ 6 & 2 & 1 \\ -2 & 2 & 1 \end{pmatrix}$.

(c) $\mathbf{A} = \begin{pmatrix} 0 & 0 & \alpha \\ 0 & \beta & 0 \\ \gamma & 0 & 0 \end{pmatrix}$. (d) $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$.

6.1.2. What is the volume of the parallelepiped generated by the three vectors $\mathbf{x}_1 = (3, 0, -4, 0)^T$, $\mathbf{x}_2 = (0, 2, 0, -2)^T$, and $\mathbf{x}_3 = (0, 1, 0, 1)^T$?

6.1.3. Using Gaussian elimination to reduce $\mathbf{A}$ to an upper-triangular matrix, evaluate $\det(\mathbf{A})$ for each of the following matrices.

(a) $\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 1 \\ 1 & 4 & 4 \end{pmatrix}$. (b) $\mathbf{A} = \begin{pmatrix} 1 & 3 & 5 \\ -1 & 4 & 2 \\ 3 & -2 & 4 \end{pmatrix}$.

(c) $\mathbf{A} = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 4 & 8 & 12 & -8 \\ 2 & 3 & 2 & 1 \\ -3 & -1 & 1 & -4 \end{pmatrix}$. (d) $\mathbf{A} = \begin{pmatrix} 0 & 0 & -2 & 3 \\ 1 & 0 & 1 & 2 \\ -1 & 1 & 2 & 1 \\ 0 & 2 & -3 & 0 \end{pmatrix}$.

(e) $\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$. (f) $\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 1 & \cdots & 1 \\ 1 & 1 & 3 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & n \end{pmatrix}$.

6.1.4. Use determinants to compute the rank of $\mathbf{A} = \begin{pmatrix} 1 & 3 & -2 \\ 0 & 1 & 2 \\ -1 & -1 & 6 \\ 2 & 5 & -6 \end{pmatrix}$.

6.1.5. Use determinants to find the values of α for which the following system possesses a unique solution.

$$\begin{pmatrix} 1 & \alpha & 0 \\ 0 & 1 & -1 \\ \alpha & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -3 \\ 4 \\ 7 \end{pmatrix}.$$

6.1.6. If $\mathbf{A}$ is nonsingular, explain why $\det\left(\mathbf{A}^{-1}\right) = 1/\det(\mathbf{A})$.

6.1.7. Explain why determinants are invariant under similarity transformations. That is, show $\det\left(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\right) = \det(\mathbf{A})$ for all nonsingular $\mathbf{P}$.

6.1.8. Explain why $\det(\mathbf{A}^*) = \overline{\det(\mathbf{A})}$.

6.1.9. (a) Explain why $|\det(\mathbf{Q})| = 1$ when $\mathbf{Q}$ is unitary. In particular, $\det(\mathbf{Q}) = \pm 1$ if $\mathbf{Q}$ is an orthogonal matrix.

(b) How are the singular values of $\mathbf{A} \in \mathcal{C}^{n\times n}$ related to $\det(\mathbf{A})$?

6.1.10. Prove that if $\mathbf{A}$ is $m \times n$, then $\det(\mathbf{A}^*\mathbf{A}) \geq 0$, and explain why $\det(\mathbf{A}^*\mathbf{A}) > 0$ if and only if $rank(\mathbf{A}) = n$.

6.1.11. If $\mathbf{A}$ is $n \times n$, explain why $\det(\alpha\mathbf{A}) = \alpha^n \det(\mathbf{A})$ for all scalars α.

6.1.12. If $\mathbf{A}$ is an $n \times n$ skew-symmetric matrix, prove that $\mathbf{A}$ is singular whenever n is odd. **Hint:** Use Exercise 6.1.11.

6.1.13. How can you build random integer matrices with $\det(\mathbf{A}) = 1$?

6.1.14. If the k^{th} row of $\mathbf{A}_{n\times n}$ is written as a sum $\mathbf{A}_{k*} = \mathbf{x}^T + \mathbf{y}^T + \cdots + \mathbf{z}^T$, where $\mathbf{x}^\mathbf{T}, \mathbf{y}^\mathbf{T}, \ldots, \mathbf{z}^\mathbf{T}$ are row vectors, explain why

$$\det(\mathbf{A}) = \det\begin{pmatrix} \mathbf{A}_{1*} \\ \vdots \\ \mathbf{x}^T \\ \vdots \\ \mathbf{A}_{n*} \end{pmatrix} + \det\begin{pmatrix} \mathbf{A}_{1*} \\ \vdots \\ \mathbf{y}^T \\ \vdots \\ \mathbf{A}_{n*} \end{pmatrix} + \cdots + \det\begin{pmatrix} \mathbf{A}_{1*} \\ \vdots \\ \mathbf{z}^T \\ \vdots \\ \mathbf{A}_{n*} \end{pmatrix}.$$

6.1.15. The CBS inequality (p. 272) says that $|\mathbf{x}^*\mathbf{y}| \leq \|\mathbf{x}\|_2^2 \|\mathbf{y}\|_2^2$ for vectors $\mathbf{x}, \mathbf{y} \in \mathcal{C}^{n\times 1}$. Use Exercise 6.1.10 to give an alternate proof of the CBS inequality along with an alternate explanation of why equality holds if and only if $\mathbf{y}$ is a scalar multiple of $\mathbf{x}$.

6.1.16. Determinant Formula for Pivots. Let $\mathbf{A}_k$ be the $k \times k$ leading principal submatrix of $\mathbf{A}_{n\times n}$ (p. 148). Prove that if $\mathbf{A}$ has an LU factorization $\mathbf{A} = \mathbf{LU}$, then $\det(\mathbf{A}_k) = u_{11}u_{22}\cdots u_{kk}$, and deduce that the k^{th} pivot is $u_{kk} = \begin{cases} \det(\mathbf{A}_1) = a_{11} & \text{for } k = 1, \\ \det(\mathbf{A}_k)/\det(\mathbf{A}_{k-1}) & \text{for } k = 2, 3, \ldots, n. \end{cases}$

6.1.17. Prove that if $rank(\mathbf{A}_{m\times n}) = n$, then $\mathbf{A}^T\mathbf{A}$ has an LU factorization with positive pivots—i.e., $\mathbf{A}^T\mathbf{A}$ is *positive definite* (pp. 154 and 559).

6.1.18. Let $\mathbf{A}(x) = \begin{pmatrix} 2-x & 3 & 4 \\ 0 & 4-x & -5 \\ 1 & -1 & 3-x \end{pmatrix}$.

(a) First evaluate $\det(\mathbf{A})$, and then compute $d\big(\det(\mathbf{A})\big)/dx$.

(b) Use formula (6.1.19) to evaluate $d\big(\det(\mathbf{A})\big)/dx$.

6.1.19. When the entries of $\mathbf{A} = [a_{ij}(x)]$ are differentiable functions of x, we define $d\,\mathbf{A}/dx = [d\,a_{ij}/dx]$ (the matrix of derivatives). For square matrices, is it always the case that $d\big(\det(\mathbf{A})\big)/dx = \det(d\mathbf{A}/dx)$?

6.1.20. For a set of functions $\mathcal{S} = \{f_1(x), f_2(x), \ldots, f_n(x)\}$ that are $n-1$ times differentiable, the determinant

$$w(x) = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}$$

is called the ***Wronskian*** of $\mathcal{S}$. If $\mathcal{S}$ is a linearly dependent set, explain why $w(x) = 0$ for every value of x. **Hint:** Recall Example 4.3.6 (p. 189).

6.1.21. Consider evaluating an $n \times n$ determinant from the definition (6.1.1).

(a) How many multiplications are required?

(b) Assuming a computer will do 1,000,000 multiplications per second, and neglecting all other operations, what is the largest order determinant that can be evaluated in one hour?

(c) Under the same conditions of part (b), how long will it take to evaluate the determinant of a 100×100 matrix?
Hint: $100! \approx 9.33 \times 10^{157}$.

(d) If all other operations are neglected, how many multiplications per second must a computer perform if the task of evaluating the determinant of a 100×100 matrix is to be completed in 100 years?