MATH 146 Linear Algebra 1, Lecture Notes by Stephen New

Chapter 4. Determinants

Permutations

4.1 Definition: A **group** is a set G together with an element $e \in G$, called the **identity** element, and a binary operation $* : G \times G \to G$, where for $a, b \in G$ we write $*(a,b)$ as $a * b$ or often simply as ab, such that

(1) $*$ is associative: $(ab)c = a(bc)$ for all $a, b, c \in G$,
(2) e is an identity: $ae = a = ea$ for all $a \in G$, and
(3) every $a \in G$ has an inverse: for every $a \in G$ there exists $b \in G$ with $ab = e = ba$.

A group G is called **abelian** when
(4) $*$ is commutative: $ab = ba$ for all $a, b \in G$.

4.2 Note: Let G be a group. Note that the identity element $e \in G$ is the unique element that satisfies Axiom (2) in the above definition because if $u \in G$ has the property that $ua = a = au$ for all $a \in G$, then in the case that $a = e$ we obtain $e = ue = e$. Also note given $a \in G$ the element b which satisfies Axiom (3) above is unique because if $ab = e$ and $ca = e$ then we have $b = eb = (ca)b = c(ab) = ce = e$.

4.3 Definition: Let G be a group. Given $a \in G$, the unique element $b \in G$ such that $ab = e = ba$ is called the **inverse** of a and is denoted by a^{-1} (unless the operation in G is addition denoted by $+$, in which case the inverse of a is also called the **negative** of a and is denoted by $-a$). We write $a^0 = e$ and for $k \in \mathbf{Z}^+$ we write $a^k = aa\cdots a$ (where the product involves k copies of a) and $a^{-k} = (a^k)^{-1}$.

4.4 Note: In a group G, we have the cancellation property: for all $a, b, c \in G$, if $ab = ac$ (or if $ca = ba$) then $b = c$. Indeed, if $ab = ac$ then

$$b = eb = (a^{-1}a)b = a^{-1}(ab) = a^{-1}(ac) = (a^{-1}a)c = ec = c.$$

4.5 Example: If R is a ring under addition and multiplication then R is also an abelian group under addition. The identity element is 0 and the inverse of $a \in R$ is $-a$. For example $\mathbf{Z}_n$, $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{C}$ are all abelian groups under addition.

4.6 Example: If R is a ring under addition and multiplication then the set

$$R^* = \big\{ a \in R \big| a \text{ is invertible} \big\}$$

is a group under multiplication. The identity element is 1 and the inverse of $a \in R$ is a^{-1}. For example $\mathbf{Z}^* = \{1, -1\}$, $\mathbf{Q}^* = \mathbf{Q} \setminus \{0\}$, $\mathbf{R}^* = \mathbf{R} \setminus \{0\}$ and $\mathbf{C}^* = \mathbf{C} \setminus \{0\}$ are all abelian groups under multiplication. For $n \in \mathbf{Z}^+$, the **group of units modulo** n is the group

$$U_n = {\mathbf{Z}_n}^* = \big\{ a \in \mathbf{Z}_n \big| \gcd(a, n) = 1 \big\}.$$

The group of units U_n is an abelian group under multiplication modulo n. When R is a ring (usually commutative), the **general linear group** $GL_n(R)$ is the group

$$GL_n(R) = \big\{ A \in M_n(R) \big| A \text{ is invertible} \big\}.$$

When $n \geq 2$, the general linear group $GL_n(R)$ is a non-abelian group under matrix multiplication.

4.7 Definition: Let X be a set. The **group of permutations** of X is the group

$$\text{Perm}(X) = \big\{ f : X \to X \big| f \text{ is bijective} \big\}$$

under composition. The identity element is the identity map $I : X \to X$ given by $I(x) = x$ for all $x \in X$. For $n \in \mathbf{Z}^+$, the n^{th} **symmetric group** is the group

$$S_n = \text{Perm}\big(\{1, 2, \cdots, n\}\big).$$

4.8 Definition: When $a_1, a_2, \cdots, a_l$ are distinct elements in $\{1, 2, \cdots, n\}$ we write

$$\alpha = (a_1, a_2, \cdots, a_l)$$

for the permutation $\alpha \in S_n$ given by

$$\begin{aligned} &\alpha(a_1) = a_2 \ , \ \alpha(a_2) = a_3 \ , \ \cdots \ , \ \alpha(a_{l-1}) = a_l \ , \ \alpha(a_l) = a_1 \\ &\alpha(k) = k \text{ for all } k \notin \{a_1, a_2, \cdots, a_l\} \, . \end{aligned}$$

Such a permutation is called a **cycle of length** l or an l**-cycle**.

4.9 Note: We make several remarks.

(1) We have $e = (1) = (2) = \cdots = (n)$.
(2) We have $(a_1, a_2, \cdots, a_l) = (a_2, a_3, \cdots, a_l, a_1) = (a_3, a_4, \cdots, a_l, a_1, a_2) = \cdots$.
(3) An l-cycle with $l \geq 2$ can be expressed *uniquely* in the form $\alpha = (a_1, a_2, \cdots, a_l)$ with $a_1 = \min\{a_1, a_2, \cdots, a_l\}$.
(4) If $\alpha = (a_1, a_2, \cdots, a_l)$ then $\alpha^{-1} = (a_l, a_{l-1}, \cdots, a_2, a_1) = (a_1, a_l, \cdots, a_3, a_2)$.
(5) If $n \geq 3$ then we have $(12)(23) = (123)$ and $(23)(12) = (132)$ so S_n is not abelian.

4.10 Definition: In S_n, given cycles α_i with $\alpha_i = (a_{i,1}, a_{i,2}, \cdots, a_{i,l_i})$, we say that the cycles α_i are **disjoint** when all the elements $a_{i,j} \in \{1, 2, \cdots, n\}$ are distinct.

4.11 Theorem: *(Cycle Notation) Every $\alpha \in S_n$ can be written as a product of disjoint cycles. Indeed every $\alpha \neq e$ can be written uniquely in the form*

$$\alpha = (a_{1,1}, \cdots, a_{1,l_1})(a_{2,1}, \cdots, a_{2,l_2}) \cdots (a_{m,1}, \cdots, a_{m,l_m})$$

with $m \geq 1$, each $l_i \geq 2$, the elements $a_{i,j}$ all distinct, each $a_{i,1} = \min\{a_{i,1}, a_{i,2}, \cdots, a_{i,l_i}\}$ and $a_{1,1} < a_{2,1} < \cdots < a_{m,1}$.

Proof: Let $e \neq \alpha \in S_n$ where $n \geq 2$. To write α in the given form, we must take $a_{1,1}$ to be the smallest element $k \in \{1, 2, \cdots, n\}$ with $\alpha(k) \neq k$. Then we must have $a_{1,2} = \alpha(a_{1,1})$, $a_{1,3} = \alpha(a_{1,2}) = \alpha^2(a_{1,1})$, and so on. Eventually we must reach l_1 such that $a_{1,1} = \alpha^{l_1}(a_{1,1})$, indeed since $\{1, 2, \cdots, n\}$ is finite, eventually we find $\alpha^i(a_{1,1}) = \alpha^j(a_{1,1})$ for some $1 \leq i < j$ and then $a_{1,1} = \alpha^{-i}\alpha^i(a_{1,1}) = \alpha^{-i}\alpha^j(a_{1,1}) = \alpha^{j-i}(a_{1,1})$. For the smallest such l_1 the elements $a_{1,1}, \cdots, a_{1,l_1}$ will be disjoint since if we had $a_{1,i} = a_{1,j}$ for some $1 \leq i < j \leq l_1$ then, as above, we would have $\alpha^{j-i}(a_{11}) = a_{11}$ with $1 \leq j - i < l_1$. This gives us the first cycle $\alpha_1 = (a_{1,1}, a_{1,2}, \cdots, a_{1,l_1})$.

If we have $\alpha = \alpha_1$ we are done. Otherwise there must be some $k \in \{1, 2, \cdots, n\}$ with $k \notin \{a_{1,1}, a_{1,2}, \cdots, a_{1,l_1}\}$ such that $\alpha(k) \neq k$, and we must choose $a_{2,1}$ to be the smallest such k. As above we obtain the second cycle $\alpha_2 = (a_{2,1}, a_{2,2}, \cdots, a_{2,l_2})$. Note that α_2 must be disjoint from α_1 because if we had $\alpha^i(a_{2,1}) = \alpha^j(a_{1,1})$ for some i, j then we would have $a_{2,1} = \alpha^{-i}\alpha^i(a_{2,1}) = \alpha^{-i}\alpha^j(a_{1,1}) = \alpha^{j-i}(a_{1,1}) \in \{a_{1,1}, \cdots, a_{1,l_1}\}$.

At this stage, if $\alpha = \alpha_1\alpha_2$ we are done, and otherwise we continue the procedure.

4.12 Definition: When a permutation $e \neq \alpha \in S_n$ is written in the unique form of the above theorem, we say that α is written in **cycle notation**. We usually write e as $e = (1)$.

4.13 Theorem: *(Even and Odd Permutations) In S_n, with $n \geq 2$,*
(1) every $\alpha \in S_n$ is a product of 2-cycles,
(2) if $e = (a_1, b_1)(a_2, b_2) \cdots (a_l, b_l)$ then l is even, that is $l = 0 \bmod 2$, and
(3) if $\alpha = (a_1, b_1)(a_2, b_2) \cdots (a_l, b_l) = (c_1, d_1)(c_2, d_2) \cdots (c_m, d_m)$ then $l = m \bmod 2$.

Solution: To prove part (1), note that given $\alpha \in S_n$ we can write α as a product of cycles, and we have

$$(a_1, a_2, \cdots, a_l) = (a_1, a_l)(a_1, a_{l-1}) \cdots (a_1, a_2)\,.$$

We shall prove part (2) by induction. First note that we cannot write e as a single 2-cycle, but we can write e as a product of two 2-cycles, for example $e = (1,2)(1,2)$. Fix $l \geq 3$ and suppose, inductively, that for all $k < l$, if we can write e as a product of k 2-cycles the k must be even. Suppose that e can be written as a product of l 2-cycles, say $e = (a_1, b_1)(a_2, b_2) \cdots (a_l, b_l)$. Let $a = a_1$. Of all the ways we can write e as a product of l 2-cycles, in the form $e = (x_1, y_1)(x_2, y_2) \cdots (x_l, y_l)$, with $x_i = a$ for some i, choose one way, say $e = (r_1, s_1)(r_2, s_2) \cdots (r_l, s_l)$ with $r_m = a$ and $r_i, s_i \neq a$ for all $i < m$, with m being as large as possible. Note that $m \neq l$ since for $\alpha = (r_1, s_1) \cdots (r_l, s_l)$ with $r_l = a$ and $r_i, s_i \neq a$ for $i < l$ we have $\alpha(s_l) = a \neq s_l$ and so $\alpha \neq e$. Consider the product $(r_m, s_m)(r_{m+1}, s_{m+1})$. This product must be (after possibly interchanging r_{m+1} and s_{m+1}) of one of the forms

$$(a,b)(a,b)\ ,\ (a,b)(a,c)\ ,\ (a,b)(b,c)\ ,\ (a,b)(c,d)$$

where a, b, c, d are distinct. Note that

$$\begin{aligned}(a,b)(a,c) &= (a,c,b) = (b,c)(a,b),\\ (a,b)(b,c) &= (a,b,c) = (b,c)(a,c),\ \text{and}\\ (a,b)(c,d) &= (c,d)(a,b)\,,\end{aligned}$$

and so in each of these three cases we could rewrite e as a product of l 2-cycles with the first occurrence of a being farther to the right, contradicting the fact that we chose m to be as large as possible. Thus the product $(r_m, s_m)(r_{m+1}, s_{m+1})$ is of the form $(a,b)(a,b)$. By cancelling these two terms, we can write e as a product of $(l-2)$ 2-cycles. By the induction hypothesis, $(l-2)$ is even, and so l is even.

Finally, to prove part (3), suppose that $\alpha = (a_1, b_1) \cdots (a_l, b_l) = (c_1, d_1) \cdots (c_m, d_m)$. Then we have

$$e = \alpha\,\alpha^{-1} = (a_1, b_1) \cdots (a_l, b_l)(c_m, d_m) \cdots (c_1, d_1).$$

By part (2), $l + m$ is even, and so $l = m \bmod 2$.

4.14 Definition: For $n \geq 2$, a permutation $\alpha \in S_n$ is called **even** if it can be written as a product of an even number of 2-cycles. Otherwise α can be written as a product of an odd number of 2-cycles, and then it is called **odd**. We define the **sign** (or the **parity**) of $\alpha \in S_n$ to be

$$(-1)^\alpha = \begin{cases} \ \ 1 \text{ if } \alpha \text{ is even,} \\ -1 \text{ if } \alpha \text{ is odd.} \end{cases}$$

4.15 Note: Note that $(-1)^e = 1$ and that for $\alpha, \beta \in S_n$, we have $(-1)^{\alpha\beta} = (-1)^\alpha (-1)^\beta$ and $(-1)^{\alpha^{-1}} = (-1)^\alpha$. Also note that when α is an l-cycle we have $(-1)^\alpha = (-1)^{l-1}$ because $(a_1, a_2, \cdots, a_l) = (a_1, a_2)(a_2, a_3) \cdots (a_{l-1}, a_l)$.

Multilinear Maps

4.16 Notation: Let R be a commutative ring. For positive integers $n_1, n_2, \cdots, n_k$, let

$$\prod_{i=1}^{k} R^{n_i} = \{(u_1, u_2, \cdots, u_k) \big| \text{each } u_i \in R^{n_i}\}.$$

Note that

$$M_{n\times k}(R) = \prod_{i=1}^{k} R^n = \{(u_1, u_2, \cdots, u_k) \big| \text{each } u_i \in R^n\}.$$

4.17 Definition: For a map $L : \prod_{i=1}^{k} R^{n_i} \to R^m$, we say that L is k-**linear** when for each index $j \in \{1, 2, \cdots, k\}$ and for all $u_i, v, w \in R^{n_j}$ and all $t \in R$ we have

$$\begin{aligned} L(u_1, \cdots, u_{j-1}, v + w, u_{j+1}, \cdots, u_n) &= L(u_1, \cdots, u_{j-1}, v, u_{j+1}, \cdots, u_n) \\ &\quad + L(u_1, \cdots, u_{j-1}, w, u_{j+1}, \cdots, u_n) \text{ , and} \\ L(u_1, \cdots, u_{j-1}, tv, u_{j+1}, \cdots, u_n) &= t\, L(u_1, \cdots, u_{j-1}, u_j, u_{j+1}, \cdots, u_n). \end{aligned}$$

For a k-linear map $L : M_{n\times k}(R) = \prod_{i=1}^{k} R^n \to R^m$ we say that L is **symmetric** when for each index $j \in \{1, 2, \cdots, k-1\}$ and for all $u_i, v, w \in R^n$ we have

$$L(u_1, \cdots, u_{j-1}, v, w, u_{j+2}, \cdots, u_n) = L(u_1, \cdots, u_{j-1}, w, v, u_{j+2}, \cdots, u_n)$$

or equivalently when for every permutation $\sigma \in S_k$ and all $u_i \in R^n$ we have

$$L(u_1, u_2, \cdots, u_k) = L(u_{\sigma(1)}, u_{\sigma(2)}, \cdots, u_{\sigma(n)}),$$

and we say that L is **skew-symmetric** when for each index $j \in \{1, 2, \cdots, k-1\}$ and for all $u_i, v, w \in R^n$ we have

$$L(u_1, \cdots, u_{j-1}, v, w, u_{j+2}, \cdots, u_n) = -\, L(u_1, \cdots, u_{j-1}, w, v, u_{j+2}, \cdots, u_n)$$

or equivalently when for every permutation $\sigma \in S_k$ and all $u_i \in R^n$ we have

$$L(u_1, u_2, \cdots, u_k) = (-1)^{\sigma} L(u_{\sigma(1)}, u_{\sigma(2)}, \cdots, u_{\sigma(n)}),$$

and we say that L is **alternating** when for each index $j \in \{1, 2, \cdots, k-1\}$ and for all $u_i, v \in R^n$ we have

$$L(u_1, \cdots, u_{j-1}, v, v, u_{j+2}, \cdots, u_n) = 0.$$

4.18 Example: As an exercise, show that for every matrix $A \in M_{m\times n}(R)$, the map $L : R^n \times R^m \to R$ given by $L(x, y) = y^T A\, x$ is 2-linear and, conversely, that given any 2-linear map $L : R^n \times R^m \to R$ there exists a unique matrix $A \in M_{m\times n}(R)$ such that $L(x, y) = y^T A\, x$ for all $x \in R^n$ and $y \in R^m$.

4.19 Theorem: *Let R be a commutative ring. Let $L : M_{n\times k} = \prod_{i=1}^{k} R^n \to R^m$ be k-linear. Then*

(1) if L is alternating then L is skew-symmetric,
(2) if L is alternating then for all indices $i, j \in \{1, 2, \cdots, k\}$ with $i < j$ and for all $u_i, v \in R^n$ we have $L(u_1, \cdots, u_{i-1}, v, u_{i+1}, \cdots, u_{j-1}, v, u_{j+1}, \cdots, u_n) = 0$, and
(3) if $2 \in R^*$ and L is skew-symmetric then L is alternating.

Proof: To prove Part (1), we suppose that L is alternating. Then for $j \in \{1, 2, \cdots, k-1\}$ and $u_i, v, w \in R^n$ we have

$$\begin{aligned} 0 &= L(u_1, \cdots, u_{j-1}, v+w, v+w, u_{j+2}, \cdots, u_n) \\ &= L(u_1, \cdots, v, v, \cdots, u_n) + L(u_1, \cdots, v, w, \cdots, u_n) \\ &\quad + L(u_1, \cdots, w, v, \cdots, u_n) + L(u_1, \cdots, w, w, \cdots, u_n) \\ &= L(u_1, \cdots, v, w, \cdots, u_n) + L(u_1, \cdots, w, v, \cdots, u_n) \end{aligned}$$

and so $L(u_1, \cdots, v, w, \cdots, u_n) = -L(u_1, \cdots, w, v, \cdots, u_n)$, hence L is skew-symmetric.

To prove Part (2) we again suppose that L is alternating. Then, as shown immediately above, L is also skew-symmetric and so for indices $i, j \in \{1, 2, \cdots, k\}$ with $i < j$ and for $u_i, v \in R^n$, in the case that $j > i+1$ we have

$$\begin{aligned} &L(u_1, \cdots, u_{i-1}, v, w, u_{i+2}, \cdots, u_{j-1}, v, u_{j+1}, \cdots, u_n) \\ &\qquad = -L(u_1, \cdots, u_{i-1}, v, v, u_{i+2}, \cdots, u_{j-1}, w, u_{j+1} \cdots, u_n) = 0. \end{aligned}$$

Finally, to prove Part (3), suppose that $2 \in R^*$ and that L is skew-symmetric. Then for an index $j \in \{1, 2, \cdots, k-1\}$ and for $u_i, v \in R^n$ we have

$$L(u_1, \cdots, u_{j-1}, v, v, u_{j+2}, \cdots, u_n) = -L(u_1, \cdots, u_{j-1}, v, v, u_{j+2}, \cdots, u_n)$$

and so $2\,L(u_1, \cdots, u_{j-1}, v, v, u_{j+2}, \cdots, u_n) = 0$. Since $2 \in R^*$ we can multiply both sides by 2^{-1} to get $L(u_1, \cdots, u_{j-1}, v, v, u_{j+2}, \cdots, u_k) = 0$.

4.20 Theorem: *Let R be a commutative ring. Given $c \in R$ there exists a unique alternating n-linear map $L : M_n(R) = \prod_{i=1}^{n} R^n \to R$ such that $L(I) = L(e_1, e_2, \cdots, e_n) = c$. This unique map L is given by*

$$L(A) = c \cdot \sum_{\sigma \in S_n} (-1)^\sigma A_{\sigma(1),1} A_{\sigma(2),2} \cdots A_{\sigma(n),n} \text{ , that is}$$
$$L(u_1, u_2, \cdots, u_n) = c \cdot \sum_{\sigma \in S_n} (-1)^\sigma (u_1)_{\sigma(1)} (u_2)_{\sigma(2)} \cdots (u_n)_{\sigma(n)}.$$

Proof: First we prove uniqueness. Suppose that $L : M_n(R) = \prod_{i=1}^{n} R^n \to R$ is alternating and n-linear with $L(I) = c$. Then for all $u_i \in R^n$ we have

$$\begin{aligned} L(u_1, u_2, \cdots, u_n) &= L\Big(\sum_{i_1=1}^{n} (u_1)_{i_1} e_{i_1} \,,\, \sum_{i_2=1}^{n} (u_2)_{i_2} e_{i_2} \,, \cdots ,\, \sum_{i_n=1}^{n} (u_n)_{i_n} e_{i_n} \Big) \\ &= \sum_{i_1, i_2, \cdots, i_n = 1}^{n} (u_1)_{i_1} (u_2)_{i_2} \cdots (u_n)_{i_n} L(e_{i_1}, e_{i_2}, \cdots, e_{i_n}). \end{aligned}$$

Note that because L is alternating, whenever we have $e_{i_j} = e_{i_k}$ for some $j \neq k$, we obtain $L(e_{i_1}, e_{i_2}, \cdots, e_{i_n}) = 0$, and so the only nonzero terms in the above sum occur when

$i_1, i_2, \cdots, i_n$ are distinct, so there is a permutation $\sigma \in S_n$ with $i_j = \sigma(j)$ for all j. Thus

$$\begin{aligned}
L(u_1, u_2, \cdots, u_n) &= \sum_{\sigma \in S_n} (u_1)_{\sigma(1)} (u_2)_{\sigma(2)} \cdots (u_n)_{\sigma(n)} L(e_{\sigma(1)}, e_{\sigma(2)}, \cdots, e_{\sigma(n)}) \\
&= \sum_{\sigma \in S_n} (u_1)_{\sigma(1)} (u_2)_{\sigma(2)} \cdots (u_n)_{\sigma(n)} (-1)^\sigma L(e_1, e_2, \cdots, e_n) \\
&= c \cdot \sum_{\sigma \in S_n} (-1)^\sigma (u_1)_{\sigma(1)} (u_2)_{\sigma(2)} \cdots (u_n)_{\sigma(n)}
\end{aligned}$$

This proves that there is a unique such map L and that it is given by the required formula.

To prove existence, it suffices to show that the map $L : M_n(R) = \prod_{i=1}^{n} R^n \to R$ given by the formula

$$L(u_1, u_2, \cdots, u_n) = c \cdot \sum_{\sigma \in S_n} (-1)^\sigma (u_1)_{\sigma(1)} (u_2)_{\sigma(2)} \cdots (u_n)_{\sigma(n)}.$$

is n-linear and alternating with $L(I) = c$. Note that this map L is n-linear because

$$\begin{aligned}
L(u_1, \cdots, v + w, \cdots, u_n) &= c \cdot \sum_{\sigma \in S_n} (-1)^\sigma (u_1)_{\sigma(1)} \cdots (v + w)_{\sigma(j)}, \cdots (u_n)_{\sigma(n)} \\
&= c \cdot \sum_{\sigma \in S_n} (-1)^\sigma (u_1)_{\sigma(1)} \cdots v_{\sigma(j)} \cdots (u_n)_{\sigma(n)} + c \sum_{\sigma \in S_n} (-1)^\sigma (u_1)_{\sigma(1)} \cdots w_{\sigma(j)} \cdots (u_n)_{\sigma(n)} \\
&= L(u_1, \cdots, v, \cdots, u_n) + L(u_1, \cdots, w, \cdots, u_n)
\end{aligned}$$

and similarly $L(u_1, \cdots, tv, \cdots, u_n) = t\, L(u_1, \cdots, v, \cdots, u_n)$.

Note that L is alternating because, given indices $i, j \in \{1, 2, \cdots, n\}$ with $i < j$, when $u_i = u_j = v$ we have

$$\begin{aligned}
L(u_1, \cdots, v, \cdots, v, \cdots, u_n) &= c \cdot \sum_{\sigma \in S_n} (-1)^\sigma (u_1)_{\sigma(1)} \cdots v_{\sigma(i)} \cdots v_{\sigma(j)} \cdots (u_n)_{\sigma(n)} \\
&= c \cdot \sum_{\sigma \in S_n, \sigma(i) < \sigma(j)} (-1)^\sigma (u_1)_{\sigma(1)} \cdots v_{\sigma(i)} \cdots v_{\sigma(j)} \cdots (u_n)_{\sigma(n)} \\
&\quad + c \cdot \sum_{\tau \in S_n, \tau(i) > \tau(j)} (-1)^\tau (u_1)_{\tau(1)} \cdots v_{\tau(i)} \cdots v_{\tau(j)} \cdots (u_n)_{\tau(n)}.
\end{aligned}$$

This is equal to 0 because the term in the first sum labeled by $\sigma \in S_n$ with $\sigma(i) < \sigma(j)$ can be paired with the term in the second sum labeled by $\tau = \sigma(i, j)$ (where $\sigma(i, j)$ denotes the composite of σ with the 2-cycle (i, j)), and then the sum of the two terms in each pair is equal to 0 because $(-1)^\tau = -(-1)^\sigma$.

Finally note that

$$\begin{aligned}
L(e_1, e_2, \cdots, e_n) &= c \cdot \sum_{\sigma \in S_n} (-1)^\sigma (e_1)_{\sigma(1)} (e_2)_{\sigma(2)} \cdots (e_n)_{\sigma(n)} \\
&= c \cdot \sum_{\sigma \in S_n} (-1)^\sigma \delta_{1,\sigma(1)} \delta_{2,\sigma(2)} \cdots \delta_{n,\sigma(n)} = c
\end{aligned}$$

because the only nonzero term in the sum occurs when $\sigma = e$.

The Determinant

4.21 Definition: Let R be a commutative ring. The unique alternating n-linear map $\det : M_n(R) \to R$ with $\det(I) = 1$ is called the **determinant** map. For $A \in M_n(R)$, the **determinant** of A, denoted by $|A|$ or by $\det(A)$, is given by

$$|A| = \det(A) = \sum_{\sigma \in S_n} (-1)^\sigma A_{\sigma(1),1} A_{\sigma(2),2} \cdots A_{\sigma(n),n}.$$

4.22 Example: As an exercise, find an explicit formula for the determinant of a 2×2 matrix and for the determinant of a 3×3 matrix.

4.23 Note: Given $c \in R$, according to the above theorem, the unique alternating n-linear map $L : M_n(R) \to R$ with $L(I) = c$ is given by $L(A) = c\,|A|$.

4.24 Theorem: *Let R be a commutative ring and let $A, B \in M_n(R)$. Then*
(1) $|A^T| = |A|$, and
(2) $|AB| = |A|\,|B|$.

Proof: To prove Part (1) note that

$$\begin{aligned}
|A| &= \sum_{\sigma \in S_n} (-1)^\sigma A_{\sigma(1),1} A_{\sigma(2),2} \cdots A_{\sigma(n),n} \\
&= \sum_{\sigma \in S_n} (-1)^\sigma A_{1,\sigma^{-1}(1)} A_{2,\sigma^{-1}(2)} \cdots A_{n,\sigma^{-1}(n)} \\
&= \sum_{\tau \in S_n} (-1)^\tau A_{1,\tau(1)} A_{2,\tau(2)} \cdots A_{n,\tau(n)} \\
&= \sum_{\tau \in S_n} (-1)^\tau (A^T)_{\tau(1),1} (A^T)_{\tau(2),2} \cdots (A^T)_{\tau(n),n} = |A^T|.
\end{aligned}$$

To prove Part (2), fix a matrix $A \in M_n(R)$ and define $L : M_r(R) \to R$ by $L(B) = |AB|$. Note that L is n-linear because

$$\begin{aligned}
L(u_1, \cdots, v + w, \cdots, u_n) &= \big|A(u_1, \cdots, v + w, \cdots, u_n)\big| \\
&= \big|(Au_1, \cdots, A(v + w), \cdots, Au_n)\big| \\
&= \big|(Au_1, \cdots, Av + Aw, \cdots, Au_n)\big| \\
&= \big|(Au_1, \cdots, Av, \cdots, Au_n)\big| + \big|(Au_1, \cdots, Aw, \cdots, Au_n)\big| \\
&= \big|A(u_1, \cdots, v, \cdots, u_n)\big| + \big|A(u_1, \cdots, w, \cdots, u_n)\big| \\
&= L(u_1, \cdots, v, \cdots, u_n) + L(u_1, \cdots, w, \cdots, u_n).
\end{aligned}$$

and similarly $L(u_1, \cdots, tv, \cdots, u_n) = t\,L(u_1, \cdots, v, \cdots, u_n)$. Note that L is alternating because

$$\begin{aligned}
L(u_1, \cdots, v, v, \cdots, u_n) &= \big|A(u_1, \cdots, v, v, \cdots, u_n)\big| \\
&= \big|(Au_1, \cdots, Av, Av, \cdots, Au_n)\big| = 0.
\end{aligned}$$

Note that $L(I) = \big|AI\big| = |A|$. Thus, by Theorem 4.20 (see Note 4.23) it follows that $L(B) = L(I)|B| = |A|\,|A|$.

4.25 Definition: Let R be a commutative ring and let $A \in M_n(R)$. We say that A is **upper triangular** when $A_{j,k} = 0$ for all $j > k$, and we say that A is **lower-triangular** when $A_{j,k} = 0$ for all $j < k$.

4.26 Theorem: *Let R be a commutative ring and let $A, B \in M_n(R)$.*

(1) If B is obtained from A by performing an elementary column operation then $|B|$ is obtained from $|A|$ as follows.
(a) if we use $C_k \leftrightarrow C_l$ with $k \neq l$ then $|B| = -|A|$,
(b) if we use $C_k \mapsto t\,C_k$ with $t \in R$ then $|B| = t\,|A|$, and
(c) if we use $C_k \mapsto C_k + tC_l$ with $t \in R$ and $k \neq l$ then $|B| = |A|$.
The same rules apply when B is obtained from A using an elementary row operation.

(2) If A is either upper-triangular or lower-triangular then $|A| = \prod_{i=1}^{n} A_{i,i}$.

Proof: If B is obtained from A using the column operation $C_k \leftrightarrow C_l$ with $k \neq l$ then $|B| = -|A|$ because the determinant map is skew-symmetric. If B is obtained from A using $C_k \mapsto t\,C_k$ with $t \in R$ then $|B| = t|A|$ because the determinant map is linear. Suppose B is obtained from A using $C_k \mapsto C_k + t\,C_l$ where $t \in R$ and $k \neq l$. Write $A = (u_1, u_2 \cdots, u_n)$ with each $u_i \in R^n$. Then since the determinant map is n-linear and alternating we have

$$\begin{aligned} |B| &= \big|(u_1, \cdots, u_k + tu_l, \cdots, u_l, \cdots, u_n)\big| \\ &= \big|(u_1, \cdots, u_k, \cdots, u_l, \cdots, u_n)\big| + t\,\big|(u_1, \cdots, u_l, \cdots, u_l, \cdots, u_n)\big| \\ &= |A| + t \cdot 0 = |A|. \end{aligned}$$

This proves Part (1) in the case of column operations. The same rules apply when using row operations because $|A^T| = |A|$.

To prove Part (2), suppose that A is upper-triangular (the case that A is lower-triangular is similar). We claim that for every $\sigma \in S_n$ with $\sigma \neq e$ we have $\sigma(i) > i$, hence $A_{\sigma(i),i} = 0$, for some $i \in \{1, 2, \cdots, n\}$. Suppose, for a contradiction, that $\sigma \neq e$ and $\sigma(i) \leq i$ for all indices i. Let k be the largest index for which $\sigma(k) < k$. Then we have $\sigma(i) = i > k$ for all $i > k$ and $\sigma(k) < k$ and $\sigma(i) \leq i < k$ for all $i < k$. This implies that there is no index i for which $\sigma(i) = k$, but this is not possible since σ is surjective. This proves the claim. Thus $|A| = \sum_{\sigma \in S_n} (-1)^\sigma A_{\sigma(1),1} A_{\sigma(2),2} \cdots A_{\sigma(n),n} = \prod_{i=1}^{n} A_{i,i}$ because the only nonzero term in the above sum occurs when $\sigma = e$.

4.27 Example: The above theorem gives us a method that we can use to calculate determinants. For example, using only row operations of the form $R_k \to R_k + t\,R_l$ we have

$$\begin{vmatrix} 1 & 3 & 2 & 4 \\ 2 & 4 & 1 & 2 \\ 3 & 5 & 4 & 1 \\ 1 & 1 & 5 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 2 & 4 \\ 0 & -2 & -3 & -6 \\ 0 & -4 & -2 & -11 \\ 0 & -2 & 3 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 2 & 4 \\ 0 & 2 & -1 & 5 \\ 0 & -4 & -2 & -11 \\ 0 & -2 & 3 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 2 & 4 \\ 0 & 2 & -1 & 5 \\ 0 & 0 & -4 & -1 \\ 0 & 0 & 2 & 4 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 3 & 2 & 4 \\ 0 & 2 & -1 & 1 \\ 0 & 0 & 2 & 11 \\ 0 & 0 & 2 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 2 & 4 \\ 0 & 2 & -1 & 5 \\ 0 & 0 & 2 & 11 \\ 0 & 0 & 0 & -7 \end{vmatrix} = -28.$$

4.28 Definition: Let R be a commutative ring and let $A \in M_n(R)$ with $n \geq 2$. We write $A^{(j,k)}$ to denote the $(n-1) \times (n-1)$ matrix which is obtained by removing the j^{th} row and the k^{th} column of A. The **cofactor matrix** of A is the matrix $\text{Cof}(A) \in M_n(R)$ with entries

$$\text{Cof}(A)_{k,l} = (-1)^{k+l} \big|A^{(l,k)}\big|.$$

4.29 Theorem: *Let R be a commutative ring and let $A \in M_n(R)$ with $n \geq 2$.*
(1) For each $k \in \{1, 2, \cdots, n\}$ we have

$$|A| = \sum_{j=1}^{n} (-1)^{j+k} A_{j,k} \big|A^{(j,k)}\big| = \sum_{j=1}^{n} (-1)^{k+j} A_{k,\,j} \big|A^{(k,\,j)}\big|.$$

(2) We have

$$A \cdot \mathrm{Cof}(A) = |A| \cdot I = \mathrm{Cof}(A) \cdot A.$$

(3) A is invertible in $M_n(R)$ if and only if $|A|$ is invertible in R, and in this case we have $|A^{-1}| = |A|^{-1}$ and

$$A^{-1} = \tfrac{1}{|A|} \mathrm{Cof}(A).$$

(4) If A is invertible then the unique solution to the equation $Ax = b$ is the element $x \in R^n$ with entries

$$x_k = \frac{|B_k|}{|A|}$$

where B_k is the matrix obtained by replacing the k^{th} column of A by b.

Proof: We have

$$\begin{aligned}
|A| &= \sum_{\sigma \in S_n} (-1)^\sigma A_{\sigma(1),1} A_{\sigma(2),2} \cdots A_{\sigma(n),n} \\
&= \sum_{j=1}^{n} \sum_{\sigma \in S_n, \sigma(k)=j} (-1)^\sigma A_{\sigma(1),1} \cdots A_{\sigma(k-1),k-1} A_{j,k} A_{\sigma(k+1),k+1} \cdots A_{\sigma(n),n} \\
&= \sum_{j=1}^{n} A_{j,k} \sum_{\sigma \in S_n, \sigma(k)=j} (-1)^\sigma {A^{(j,k)}}_{\tau(1),1} \cdots {A^{(j,k)}}_{\tau(k-1),k-1} {A^{(j,k)}}_{\tau(k),k} \cdots {A^{(j,k)}}_{\tau(k-1),k-1}
\end{aligned}$$

where $\tau = \tau(\sigma) \in S_{n-1}$ is the permutation defined as follows:

$$\text{if } i < k \quad \tau(i) = \begin{cases} \sigma(i) & \text{if } \sigma(i) < j, \\ \sigma(i) - 1 & \text{if } \sigma(i) > j, \end{cases} \quad \text{and if } i > k \quad \tau(i-1) = \begin{cases} \sigma(i) & \text{if } \sigma(i) < j, \\ \sigma(i) - 1 & \text{if } \sigma(i) > j, \end{cases}$$

or equivalently, τ is the composite

$$\tau = (n, n-1, \cdots, j+1, j)\, \sigma\, (k, k+1, \cdots, n-1, n).$$

Note that $(-1)^\tau = (-1)^{n-j}(-1)^\sigma(-1)^{n-k}$ and so we have $(-1)^\sigma = (-1)^{j+k}(-1)^\tau$. Thus

$$\begin{aligned}
|A| &= \sum_{j=1}^{n} A_{j,k} \sum_{\tau \in S_{n-1}} (-1)^{j+k} (-1)^\tau {A^{(j,k)}}_{\tau(1),1} \cdots {A^{(j,k)}}_{\tau(n-1),n-1} \\
&= \sum_{j=1}^{n} (-1)^{j+k} A_{j,k} \big|A^{(j,k)}\big|.
\end{aligned}$$

The proof that $|A| = \sum\limits_{j=1}^{n} (-1)^{k+j} A_{k,j} \big|A^{(k,j)}\big|$ is similar (or it follows from the formula $|A^T| = |A|$). This completes the proof of Part (1).

To prove Part (2) we note that

$$\Big(\text{Cof}(A)\cdot A\Big)_{k,l}=\sum_{j=1}^{n}\text{Cof}(A)_{k,j}A_{j,l}=\sum_{j=1}^{n}(-i)^{k+j}A_{j,l}\,\big|A^{(j,k)}\big|.$$

By Part (1), the sum on the right is equal to the determinant of the matrix $B^{(k,l)}\in M_n(R)$ which is obtained from A by replacing the k^{th} column of A by a copy of its l^{th} column. Since $B^{(k,l)}$ is equal to A when $k=l$, and $B^{(k,l)}$ has two equal columns when $k\neq l$ we have

$$\Big(\text{Cof}(A)\cdot A\Big)_{k,l}=\big|B^{(k,l)}\big|=\left\{\begin{array}{l}|A|\text{ if }k=l,\\ \;0\;\text{ if }k\neq l\end{array}\right\}=|A|\cdot\delta_{k,l}.$$

This proves that $\text{Cof}(A)\cdot A=|A|\cdot I$. A similar proof shows that $A\cdot\text{Cof}(A)=|A|\cdot I$.

If A is invertible in $M_n(R)$, then we have $|A|\,|A^{-1}|=|A\cdot A^{-1}|=|I|=1$ and similarly $|A^{-1}|\,|A|=1$ and so $|A|$ and $|A^{-1}|$ are invertible in R with $|A^{-1}|=|A|$. Conversely, the formulas in Part (2) show that if $|A|$ is invertible in R then A is invertible in $M_n(R)$ with $A^{-1}=\frac{1}{|A|}\text{Cof}(A)$. This proves Part (3).

Part (4) now follows from Parts (1) and (3). Indeed if A is invertible then the solution to $Ax=b$ is given by $x=A^{-1}b$ and so

$$\begin{aligned}x_k=\big(A^{-1}b\big)_k&=\tfrac{1}{|A|}\big(\text{Cof}(A)\,b\big)_k=\tfrac{1}{|A|}\sum_{j=1}^{n}\text{Cof}(A)_{k,j}b_j\\&=\tfrac{1}{|A|}\sum_{j=1}^{n}(-1)^{k+j}b_j\,\big|A^{(j,k)}\big|=\tfrac{1}{|A|}\,|B_k|\end{aligned}$$

where B_k is the matrix obtained by replacing the k^{th} column of A by b.

4.30 Definition: For a matrix $A\in M_n(R)$, the first of the two sums in Part (1) of the above theorem is called the **cofactor expansion** of $|A|$ **along the** k^{th} **column** of A, and the second sum is called the **cofactor expansion** of $|A|$ **along the** k^{th} **row** of A.

4.31 Example: Using row operations of the form $R_k\mapsto R_k+t\,R_l$, together with cofactor expansions along various columns, we have

$$\begin{aligned}\begin{vmatrix}2&1&3&4&2\\5&2&4&3&1\\1&0&2&3&1\\2&0&1&2&0\\4&3&5&6&2\end{vmatrix}&=\begin{vmatrix}2&1&3&4&2\\1&0&-2&-5&-3\\1&0&2&3&1\\2&0&1&2&0\\-2&0&-4&-6&-4\end{vmatrix}=-\begin{vmatrix}1&-2&-5&-3\\1&2&3&1\\2&1&2&0\\-2&-4&-6&-4\end{vmatrix}=-\begin{vmatrix}4&4&4&0\\1&2&3&1\\2&1&2&0\\2&4&6&0\end{vmatrix}\\&=-\begin{vmatrix}4&4&4\\2&1&2\\2&4&6\end{vmatrix}=-\begin{vmatrix}-4&0&-4\\2&1&2\\-6&0&-2\end{vmatrix}=-\begin{vmatrix}-4&-4\\-6&-2\end{vmatrix}=-\begin{vmatrix}8&0\\-6&-2\end{vmatrix}=16.\end{aligned}$$

4.32 Example: From the formula in Part (2), if $ad-bc\neq 0$ then we have

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}^{-1}=\frac{1}{ad-bc}\begin{pmatrix}d&b\\-c&a\end{pmatrix}.$$

As an exercise, find a similar formula for the inverse of a 3×3 matrix.