

Chapter IV. Determinants.

IV.1 The Permutation Group S_n .

The **permutation group** S_n consists of all bijections $\sigma : [1, n] \rightarrow [1, n]$ where $[1, n] = \{1, \dots, n\}$, with composition of operators

$$\sigma_1 \circ \sigma_2(k) = \sigma_1(\sigma_2(k)) \quad \text{for } 1 \leq k \leq n$$

as the group operation. The identity element e is the identity map $\text{id}_{[1,n]}$ such that $e(k) = k$, for all $k \in [1, n]$. We recall that a **group** is any set G equipped with a binary operation $(*)$ satisfying the following axioms:

1. ASSOCIATIVITY: $x * (y * z) = (x * y) * z$;
2. IDENTITY ELEMENT: There is an $e \in G$ such that $e * x = x = x * e$, for all $x \in G$;
3. INVERSES: Every $x \in G$ has a “two-sided inverse,” an element $x^{-1} \in G$ such that $x^{-1} * x = x * x^{-1} = e$.

We do not assume that the system $(G, *)$ is commutative, with $x * y = y * x$; a group with this extra property is a *commutative group*, also referred to as an *abelian group*. Here are some examples of familiar groups.

1. The integers $(\mathbb{Z}, +)$ become a commutative group when equipped with $(+)$ as the group operation; multiplication (\cdot) *does not* make \mathbb{Z} a group. (Why?)
2. Any vector space equipped with its $(+)$ operation is a commutative group, for instance $(\mathbb{K}^n, +)$;
3. The set $(\mathbb{C}^\times, \cdot) = \mathbb{C} \setminus \{0\}$ of nonzero complex numbers equipped with complex multiplication (\cdot) is a commutative group; so is the subset $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ (unit circle in the complex plane) because $|z|, |w| = 1 \Rightarrow |zw| = |z| \cdot |w| = 1$ and $|1/z| = 1/|z| = 1$.
4. GENERAL LINEAR GROUP. The set $\text{GL}(n, \mathbb{K}) = \{A \in \text{M}(n, \mathbb{K}) : \det(A) \neq 0\}$ of invertible $n \times n$ matrices is a group when equipped with matrix multiply as the group operation. It is noncommutative when $n \geq 2$. Validity of the group axioms for (GL, \cdot) follows because

$$\det(AB) = \det(A) \cdot \det(B) \quad \det(I) = 1 \quad \det(A^{-1}) = \frac{1}{\det(A)} \quad ,$$

and a matrix A has a two-sided inverse $\Leftrightarrow \det(A) \neq 0$.

SPECIAL LINEAR GROUP. These properties of the determinant imply that the subset $\text{SL}(n, \mathbb{K}) = \{A \in \text{M}(n, \mathbb{K}) : \det(A) = 1\}$ equipped with matrix multiply is also a (noncommutative) group;

5. The set of permutations $(\text{Per}(X), \circ)$, the bijections on a set X of n distinct objects, is also a group when equipped with composition (\circ) as its product operation. No matter what the nature of the objects being permuted, we can restrict attention to permutations of the set of integers $[1, n]$ by labeling the original objects, and then we have the group S_n .

Permutations. The simplest permutations are the k -cycles.

1.1. Definition. An ordered list (i_1, \dots, i_k) of k distinct indices in $[1, n] = \{1, \dots, n\}$ determines a k -cycle in S_n , the permutation that acts in the following way on the set $X = [1, n]$.

$$(29) \quad \sigma \text{ maps } \begin{cases} i_1 \rightarrow i_2 \rightarrow \dots \rightarrow i_k \rightarrow i_1 & \text{(a one-step "cyclic shift" of list entries)} \\ j \rightarrow j & \text{for all } j \text{ not in the list } \{i_1, \dots, i_k\} \end{cases}$$

A 1-cycle (k) is just the identity map id_X so we seldom indicate them explicitly, though it is permissible and sometimes quite useful to do so. The **support** of a k -cycle is the set of entries $\text{supp}(\sigma) = \{i_1, \dots, i_k\}$, in no particular order. The support of a one-cycle (k) is the one-point set $\{k\}$.

The order of the entries in the symbol $\sigma = (i_1, \dots, i_k)$ matters, but cycle notation is ambiguous: k different symbols

$$(i_1, \dots, i_k) = (i_2, \dots, i_k, i_1) = (i_3, \dots, i_k, i_1, i_2) = \dots = (i_k, i_1, \dots, i_{k-1})$$

obtained by "cyclic shifts" of the list entries in σ ; all describe the same operation in S_n . Thus a k -cycle might best be described by a "cyclic list" of the sort shown below, rather than a linearly ordered list, but such diagrams are a bit cumbersome for the printed page. If we change the cyclic order of the indices we get a new operator. Thus $(1, 2, 3) = (2, 3, 1) = (3, 1, 2) \neq (1, 3, 2)$ because $(1, 2, 3)$ sends $1 \rightarrow 2$ while $(1, 3, 2)$ sends $1 \rightarrow 3$.

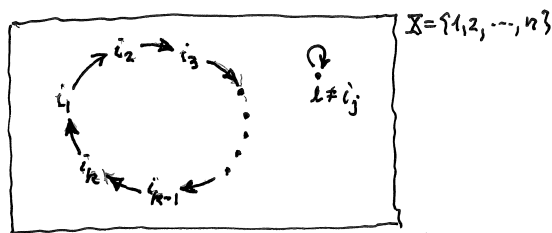


Figure 4.1. Action of the k -cycle $\sigma = (i_1, \dots, i_k)$ on $X = \{1, 2, \dots, n\}$. Points ℓ not in the "support set" $\text{supp}(\sigma) = \{i_1, \dots, i_k\}$ remain fixed; those in the support set are shifted one step clockwise in this cyclically ordered list. This σ is a "1-shift." (A 2-shift would move points 2 steps in the cyclic order, sending $i_1 \rightarrow i_3$ to ... etc.

One (cumbersome) way to describe general elements $\sigma \in S_n$ employs a data array to show where each $k \in [1, n]$ ends up:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ j_1 & j_2 & j_3 & \dots & j_n \end{pmatrix}$$

More efficient notation is afforded by the fact that every permutation σ can be uniquely written as a product of cycles with *disjoint* supports, which means that the factors commute.

1.2. Exercise. If $\sigma = (i_1, \dots, i_r)$, $\tau = (j_1, \dots, j_s)$ act on *disjoint* sets of indices, show that these operators commute. This is no longer true if the sets of indices overlap. Check this by computing the effect of the following products $\sigma\tau(k) = \sigma(\tau(k))$ of permutations in S_5 .

1. $(1, 2, 3)(2, 4);$

2. $(2, 4)(1, 2, 3)$.

Is either product a cycle?

Thus the order of factors in a product of cycles is irrelevant if the cycles are disjoint.

The product of two cycles $\sigma\tau = \sigma \circ \tau$ is a composition of operators, so the action of $\sigma\tau = \sigma \circ \tau$ on an element $k \in [1, n]$ is evaluated by feeding k into the product *from the right* as below. Taking $\sigma = (1, 2)$, and $\tau = (1, 2, 3)$ in S_5 we have

$$\sigma\tau : k \xrightarrow{(1,2,3)} (1, 2, 3) \cdot k \xrightarrow{(1,2)} (1, 2) \cdot ((1, 2, 3) \cdot k) = ((1, 2)(1, 2, 3) \cdot k)$$

To determine the net effect we track what happens to each k :

Action			Net Effect
1	$\xrightarrow{(1,2,3)}$	2	$\xrightarrow{(1,2,3)}$ 1
2	\longrightarrow	3	\longrightarrow 3
3	\longrightarrow	1	\longrightarrow 2
4	\longrightarrow	4	\longrightarrow 4
5	\longrightarrow	5	\longrightarrow 5

Thus the product $(1, 2)(1, 2, 3)$ is equal to $(2, 3) = (1)(2, 3)(4)(5)$, when we include redundant 1-cycles. On the other hand $(1, 2, 3)(1, 2) = (1, 3)$ which shows that cycles need not commute if their supports overlap. As another example we have

$$(1, 2, 3, 4)^2 = (1, 3)(2, 4)$$

which shows that a power σ^k of a cycle need not be a cycle, although it is a product of disjoint cycles. We cite without proof the fundamental cycle decomposition theorem.

1.3. Theorem (Cycle Decomposition of Permutations). *Every $\sigma \in S_n$ is a product of disjoint cycles. This decomposition is unique (up to order of the commuting factors) if we include the 1-cycles needed to account for all indices $k \in [1, n]$.*

1.4. Exercise. Write

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 6 & 5 & 1 & 3 \end{pmatrix}$$

as a product of disjoint commuting cycles.

Hint: Start by tracking $1 \rightarrow 2 \rightarrow 4 \rightarrow \dots$ until a cycle is completed; then feed σ the first integer not included in the previous cycle, etc.

1.5. Exercise. Evaluate the net action of the following products of cycles

1. $(1, 2)(1, 3)$ in S_3 ;
2. $(1, 2)(1, 3)$ in S_6 ;
3. $(1, 2)(1, 2, 3, 4, 5)$ in S_5 ;
4. $(1, 2, 3, 4, 5)(1, 2)$ in S_5 ;
5. $(1, 2)^2$ in S_5 ;
6. $(1, 2, 3)^2$ in S_5 .

Write each as a product of disjoint cycles.

1.6. Exercise. Determine the inverses σ^{-1} of the following elements in S_5

1. $(1, 2)$;
2. $(1, 2, 3)$;
3. Any 2-cycle (i_1, i_2) with $i_1 \neq i_2$;
4. Any k -cycle (i_1, \dots, i_k) .

1.7. Exercise. Evaluate the following products in S_n as products of disjoint cycles

1. $(1, 5)(1, 4)(1, 3)(1, 2)$;

2. $(1, 2)(1, 3)(1, 4)(1, 5);$
3. $(1, k)(1, 2, \dots, k-1).$

1.8. Exercise. The **order** $o(\sigma)$ of a permutation σ is the smallest integer $m \geq 1$ such that $\sigma^m = \sigma \cdots \sigma = e$.

1. Prove that every k -cycle has order $o(\sigma) = k$.
2. Verify that the r^{th} power σ^r of a k -cycle $\sigma = (i_1, \dots, i_k)$ is an “ r -shift” that moves every entry clockwise r steps in the cyclically ordered list of Figure 4.1.
3. If σ is a 6-cycle its square $\sigma^2 = \sigma \circ \sigma$ is a cyclic 2-shift of the entries (i_1, \dots, i_6) . What is the order of this element in S_n ?

Hint: By relabeling, it suffices to consider the standard 6-cycle $(1, 2, 3, 4, 5, 6)$ in answering (3.)

The only element in S_n of order 1 is the identity e ; two-cycles have order 2. As noted above, in (2.) the powers σ^r of a k -cycle need not be cycles (but sometimes they are).

Parity of a Permutation. In a different direction we note that the 2-cycles (i, j) generate the entire group S_n in the sense that every $\sigma \in S_n$ can be written as a product $\sigma = \tau_1 \cdots \tau_r$ of 2-cycles. However these factors are not necessarily disjoint and need not commute, and such decompositions are far from unique since we have, for example,

$$e = (1, 2)^2 = (1, 2)^4 = (1, 3)^2 \text{ etc..}$$

Nevertheless an important aspect of such factorizations is unique, namely its **parity**

$$\text{sgn}(\sigma) = (-1)^r$$

where $r = \#(2\text{-cycles in the factorization } \sigma = \tau_1, \dots, \tau_r)$. That means the elements $\sigma \in S_n$ fall into two disjoint classes: *even permutations* that can be written as a product of an even number of 2-cycles, and *odd permutations*. It is not obvious that all 2-cycle decompositions of a given permutation have the same parity. We prove that next, and then show how to compute $\text{sgn}(\sigma)$ effectively.

We first observe that a decomposition into 2-cycles always exists. By Theorem 1.3 it suffices to show that any k -cycle can be so decomposed. For 1-cycles this is obvious since $(k) = e = (1, 2) \cdot (1, 2)$. When $k > 1$ it is easy to check that

$$(1, 2, \dots, k) = (1, k) \cdot \dots \cdot (1, 3)(1, 2)$$

(with $k-1$ factors)

1.9. Exercise. Verify the preceding factorization of the cycle $(1, 2, \dots, k)$. Then by relabeling deduce that $(i_1, \dots, i_k) = (i_1, i_k)(i_1, i_{k-1}) \cdots (i_1, i_2)$ for any k -cycle.

Note: This is an example of “proof by relabeling.”

Once we verify that the parity is well defined, this tell us how to recognize the parity of any k -cycle

$$(30) \quad \text{sgn}(i_1, i_2, \dots, i_k) = (-1)^{k-1} \quad \text{for all } k > 0$$

1.10. Theorem (Parity). All decompositions $\sigma = \tau_1 \cdots \tau_r$ of a permutation as a product of 2-cycles have the same parity $\text{sgn}(\sigma) = (-1)^r$.

Proof: The group S_n acts on the space of polynomials $\mathbb{K}[\mathbf{x}] = \mathbb{K}[x_1, \dots, x_n]$ by permuting the variables

$$(\sigma \cdot f)(x_1, \dots, x_n) = f(x_{\sigma(1)}, \dots, x_{\sigma(n)})$$

For instance $(1, 2, 3) \cdot f(x_1, x_2, x_3, x_4, x_5) = f(x_2, x_3, x_1, x_4, x_5)$. This is a “covariant group action” in the sense that

$$(\sigma\tau) \cdot f = \sigma \cdot (\tau \cdot f) \quad \text{and} \quad e \cdot f = f$$

for all f and all $\sigma, \tau \in S_n$. The notation makes this a bit tricky to prove; one way to convince yourself is to write

$$\begin{aligned} \sigma \cdot (\tau \cdot f)(x_1, \dots, x_n) &= \tau \cdot f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) \\ &= \tau \cdot f(w_1, \dots, w_n) \Big|_{w_1=x_{\sigma(1)}, \dots, w_n=x_{\sigma(n)}} \\ &= f(w_{\tau(1)}, \dots, w_{\tau(n)}) \Big|_{w_k=x_{\sigma(k)}} \\ &= f(x_{\sigma(\tau(1))}, \dots, x_{\sigma(\tau(n))}) \\ &= f(x_{(\sigma\tau)(1)}, \dots, x_{(\sigma\tau)(n)}) = (\sigma\tau) \cdot f(x_1, \dots, x_n) \end{aligned}$$

Now consider the polynomial in n unknowns $\phi \in \mathbb{K}[x_1, \dots, x_n]$ given by

$$\phi(x_1, \dots, x_n) = \prod_{i < j} (x_i - x_j).$$

We claim that $\sigma \cdot \phi = (-1)^{\text{sgn}(\sigma)} \phi$ for any permutation σ ; by “covariance” it follows that $\sigma \cdot \phi = (-1)^r \phi$ if σ is a product $\tau_1 \cdot \dots \cdot \tau_r$ of r two-cycles. Since the definition of $\sigma \cdot \phi$ makes no reference to 2-cycle decompositions we will conclude that $(-1)^r$ must be the same for all such decompositions of σ , completing the proof.

To show that $\tau \cdot \phi = (-1)\phi$ for a 2-cycle (i, j) we may assume $i < j$. Note that the terms $x_k - x_\ell$ ($k < \ell$) not involving i or j are unaffected when we switch $x_i \leftrightarrow x_j$. The remaining terms are of three types.

CASE 1: Terms involving both i and j . The only such term is $x_i - x_j$ which becomes

$$\sigma \cdot (x_i - x_j) = x_j - x_i = (-1)(x_i - x_j),$$

suffering a change of sign.

CASE 2: Terms involving only i . The possibilities (for $k \neq j$) are listed below

Terms	$x_k - x_i$ $1 \leq k < i$	$x_i - x_k$ $i < k < j$	$x_i - x_k$ $j < k \leq n$
#(Terms)	$i - 1$	$j - i - 1$	$n - j$
Effect of $x_i \leftrightarrow x_j$ on sign of term	No change (since $i < j < k$)	$x_i - x_k \rightarrow x_j - x_k$ $= (-1)(x_k - x_j)$	No change (since $i < j < k$)

CASE 3: Terms involving only j . These are (for $k \neq i$).

Terms	$x_k - x_j$ $1 \leq k < i$	$x_k - x_j$ $i < k < j$	$x_j - x_k$ $j < k \leq n$
#(Terms)	$j - 1$	$j - i - 1$	$n - j$
Effect of $x_i \leftrightarrow x_j$ on sign of term	No change (since $i < j < k$)	$x_k - x_j \rightarrow x_k - x_i$ $= (-1)(x_i - x_k)$	No change

The effect of switching $x_i \leftrightarrow x_j$ is to permute the terms in $\prod_{k < l} (x_k - x_l)$ changing the sign of some, so the product gets multiplied by +1 or -1. Counting the number of sign changes in all cases we see that

$$(-1)^{\#(\text{changes})} = (-1)^{1+\text{even}} = -1$$

as claimed. \square

1.11. Corollary. *The parity map $\text{sgn} : S_n \rightarrow \{\pm 1\}$, defined by $\text{sgn}(\sigma) = (-1)^r$ if σ can be written as a product of r two-cycles, has the following algebraic properties*

1. $\text{sgn}(e) = +1$;
2. $\text{sgn}(\sigma\tau) = \text{sgn}(\sigma) \cdot \text{sgn}(\tau)$;
3. $\text{sgn}(\sigma^{-1}) = (\text{sgn}(\sigma))^{-1} = \text{sgn}(\sigma)$ (since $\text{sgn} = \pm 1$).

Proof: Obviously $\text{sgn}(e) = 1$ since we may write $e = (1, 2)^2$. If $\sigma = c_1 \cdot \dots \cdot c_r$ and $\tau = c'_1 \cdot \dots \cdot c'_s$ where c_i, c'_j are 2-cycles, then $\sigma\tau = c_1 \cdot \dots \cdot c_r c'_1 \cdot \dots \cdot c'_s$ is a product of $r + s$ cycles, proving (2.). The third property follows because

$$1 = \text{sgn}(e) = \text{sgn}(\sigma\sigma^{-1}) = \text{sgn}(\sigma) \cdot \text{sgn}(\sigma^{-1})$$

since the only values of sgn are ± 1 . \square

IV.2 Determinants.

The previous digression about the permutation group S_n is needed to formulate the natural definition of $\det(A)$ for an $n \times n$ matrix $A \in M(n, \mathbb{K})$, or of $\det(T)$ for a linear operator $T : V \rightarrow V$ on a finite dimensional vector space.

Any discussion that formulates this definition in terms of “expansion by minors” is confusing the natural definition of *det* with a commonly use *algorithm* for computing its value. Here is the real definition:

2.1. Definition. *If $A \in M(n, \mathbb{K})$, we define its **determinant** to be*

$$(31) \quad \det(A) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \cdot a_{1, \sigma(1)} \cdot \dots \cdot a_{n, \sigma(n)} = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \cdot \prod_{i=1}^n a_{i, \sigma(i)}$$

The products in this sum are obtained by taking $\sigma \in S_n$ and using it to select one entry from each row, taking each entry from a different column. Thus each σ determines a “template” for selecting matrix entries that are to be multiplied together (the product then weighted by the signature $\text{sgn}(\sigma)$ of the permutation). The idea is illustrated in Figure 4.2.

Many properties can be read directly from definition but the all-important multiplicative property $\det(AB) = \det(A) \cdot \det(B)$ is tricky no matter what definition we start from. We begin with several easy properties:

2.2. Theorem. *If $A \in M(n, \mathbb{K})$ and $c \in \mathbb{K}$ we have*

1. $\det(I_{n \times n}) = 1$;
2. $\det(cA) = c^n \cdot \det(A)$ if A is $n \times n$;
3. $\det(A^t) = \det(A)$;
4. When $\mathbb{K} = \mathbb{C}$ we have $\det(\bar{A}) = \overline{\det(A)}$ where $\bar{z} =$ the complex conjugate of $z = x + iy$.

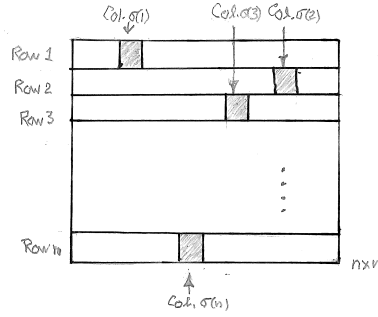


Figure 4.2. A permutation $\sigma \in S_n$ determines a “template” for selecting matrix entries by marking the address $(i, \sigma(i))$ – the one in Row i , Column $\sigma(i)$. Each row contains exactly one marked spot, and likewise for each column.

5. If A is “upper triangular,” so

$$A = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix},$$

then $\det(A) = \prod_{k=1}^n a_{kk}$ is the product of the diagonal entries.

Proof: Assertions (1.), (2.), (4.) are all trivial; we leave their proof to the reader. In (5.) the typical product $\pm a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$ in the definition of $\det(A)$ will equal 0 if any factor is zero. But unless $\sigma(k) = k$ for all k , there will be some row such that $\sigma(k) > k$ and some other row such that $\sigma(\ell) < \ell$. The resulting template includes a matrix entry below the diagonal, making the product for this template zero. The only permutation contributing a term to the sum (31) is $\sigma = e$, and that term is equal to $a_{11} \cdots a_{nn}$ as in (5.)

For (3.) we note that

$$\det(A^t) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) (b_{1,\sigma(1)} \cdots b_{n,\sigma(n)})$$

if $B = A^t = [b_{ij}]$. By definition of A^t , $b_{ij} = a_{ji}$ so the typical term becomes

$$b_{1,\sigma(1)} \cdots b_{n,\sigma(n)} = a_{\sigma(1),1} \cdots a_{\sigma(n),n}$$

However, we may write $a_{\sigma(j),j} = a_{\sigma(j),\sigma^{-1}(\sigma(j))}$ for each j , and then

$$\det(A^t) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) b_{1,\sigma(1)} \cdots b_{n,\sigma(n)} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}$$

Note that $\prod_i a_{\sigma(i),i} = \prod_{i=1}^n a_{\sigma(i),\sigma^{-1}(\sigma(i))}$, so if we replace the dummy index i in the product with $j = \sigma(i)$ the product becomes $\prod_{j=1}^n a_{j,\sigma^{-1}(j)}$ and

$$\det(A^t) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \prod_{j=1}^n a_{j,\sigma^{-1}(j)}.$$

Next, write $\tau = \sigma^{-1}$. The τ run through all of S_n as σ runs through S_n because S_n is a group. (This is our first encounter with the “group” property of S_n .) Furthermore

$\text{sgn}(\tau) = \text{sgn}(\sigma^{-1})$ so that

$$\det(A^t) = \sum_{\tau \in S_n} \text{sgn}(\tau) \cdot \prod_{j=1}^n a_{j, \tau(j)} = \det(A) \quad \square$$

The following observation will play a pivotal role in further discussion of determinants.

2.3. Lemma. *If B is obtained from A by interchanging two rows (or two columns) then*

$$\det(B) = (-1) \cdot \det(A) .$$

Proof: We do the case of column interchange. If $A = [a_{ij}]$ then $B = [b_{ij}]$ with $b_{ij} = a_{i\tau(j)}$; i.e. $\text{Col}_j(B) = \text{Col}_{\tau(j)}(A)$, for $1 \leq j \leq n$, where τ is the two-cycle $\tau = (k, \ell)$ that switches the column indices when we interchange $\text{Col}_\ell(A) \leftrightarrow \text{Col}_k(A)$. Then for any $\sigma \in S_n$, we have

$$b_{1, \sigma(1)} \cdot \dots \cdot b_{n, \sigma(n)} = a_{1, \tau\sigma(1)} \cdot \dots \cdot a_{n, \tau\sigma(n)}$$

But S_n is a *group* so $S_n\sigma = S_n$ and the elements $\tau\sigma$ run through all of S_n as τ runs through S_n ; furthermore, because τ is a 2-cycle we have $\text{sgn}(\tau) = -1$ and $\text{sgn}(\tau\sigma) = \text{sgn}(\tau)\text{sgn}(\sigma) = (-1) \cdot \text{sgn}(\sigma)$. Thus

$$\begin{aligned} \det(B) &= \sum_{\sigma \in S_n} \text{sgn}(\sigma) \cdot \prod_{i=1}^n b_{i, \sigma(i)} = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \cdot \prod_{i=1}^n a_{i, \tau\sigma(i)} \\ &= \sum_{\sigma \in S_n} \text{sgn}(\tau)^{-1} \text{sgn}(\tau\sigma) \cdot \prod_{i=1}^n a_{i, \tau\sigma(i)} \\ &= \text{sgn}(\tau) \cdot \sum_{\mu \in S_n} \text{sgn}(\mu) \cdot \prod_{i=1}^n a_{i, \mu(i)} = (-1) \cdot \det(A) \quad \square \end{aligned}$$

2.4. Exercise. Use the previous results to show that $\det(A) = 0$ if either:

1. A has two identical rows (or columns);
2. A has a row (or column) consisting entirely of zeros.

Recall the definition of the “elementary row operations” on a matrix A .

- TYPE I: $R_i \leftrightarrow R_j$: interchange Row_i and Row_j ;
- TYPE II: $R_i \rightarrow \lambda \cdot R_i$: multiply Row_i by λ ($\lambda \in \mathbb{K}$);
- TYPE III: $R_i \rightarrow R_i + \lambda R_j$: Add to Row_i any scalar multiple of a different row R_j (leaving Row_j unaltered).

The effect of the first two operations on the determinant of a square matrix is easy to evaluate.

2.5. Exercise. Prove that if B has $R_i(B) = \lambda \cdot R_i(A)$ with all other rows unchanged, then $\det(B) = \lambda \cdot \det(A)$.

To deal with Type III operations we first observe that the map $\det : M(n, \mathbb{K}) \rightarrow \mathbb{K}$ is a **multilinear function** of the rows or columns of A .

2.6. Lemma. *If the i^{th} row of a matrix A is decomposed as a linear combination $R_i = aR'_i + bR''_i$ of two other rows of the same length, then*

$$\det(A) = \det \begin{pmatrix} R_1 \\ \vdots \\ aR'_i + bR''_i \\ \vdots \\ R_n \end{pmatrix} = a \cdot \det \begin{pmatrix} R_1 \\ \vdots \\ R'_i \\ \vdots \\ R_n \end{pmatrix} + b \cdot \det \begin{pmatrix} R_1 \\ \vdots \\ R''_i \\ \vdots \\ R_n \end{pmatrix} = a \cdot \det(A') + b \cdot \det(A'')$$

In other words $\det(A)$ is a multilinear function of its rows: If we vary only R_i holding the other rows fixed, the determinant is a linear function of R_i .

Proof: If $R'_i = (x_1, \dots, x_n)$ and $R''_i = (y_1, \dots, y_n)$, then $A_{ij} = ax_j + by_j$ and

$$\begin{aligned} \det(A) &= \sum_{\sigma \in S_n} \text{sgn}(\sigma) \cdot (a_{1,\sigma(1)} \cdots (ax_{\sigma(i)} + by_{\sigma(i)}) \cdots a_{n,\sigma(n)}) \\ &= a \cdot \sum_{\sigma \in S_n} \text{sgn}(\sigma) \cdot (a_{1,\sigma(1)} \cdots x_{\sigma(i)} \cdots a_{n,\sigma(n)}) \\ &\quad + b \cdot \sum_{\sigma \in S_n} \text{sgn}(\sigma) \cdot (a_{1,\sigma(1)} \cdots y_{\sigma(i)} \cdots a_{n,\sigma(n)}) \\ &= a \cdot \det(A') + b \det(A'') \end{aligned}$$

as claimed. \square

2.7. Corollary. *If B is obtained from A by a Type III row operation $R_i \rightarrow R_i + cR_j$ ($j \neq i$) then $\text{Row}_i(B) = R_i + cR_j$ and*

$$\det(B) = \det \begin{pmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_j \\ \vdots \\ R_n \end{pmatrix} + c \cdot \det \begin{pmatrix} R_1 \\ \vdots \\ R_j \\ \vdots \\ R_j \\ \vdots \\ R_n \end{pmatrix} = \det(A) + 0 = \det(A)$$

because the second matrix has a repeated row.

Row Operations, Determinants, and Inverses. Every row operation on an $n \times m$ matrix A can be implemented by multiplying A on the left by a suitable $n \times n$ “elementary matrix” E ; the corresponding column operation is achieved by multiplying A on the right by the transpose E^t .

- TYPE I. $(\text{Row}_i) \rightarrow \lambda \cdot (\text{Row}_i)$: is equivalent to sending A to $E_I A$ where

$$E_I = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & \lambda & \\ & & & \ddots \\ 0 & & & & 1 \end{pmatrix}$$

Obviously $\det(E_I) = \lambda$ and

$$\det(E_I A) = \det(E_I) \cdot \det(A) = \lambda \cdot \det(A)$$

- TYPE II. $(Row_i) \leftrightarrow (Row_j)$: Now the result is achieved using the matrix

$$E_{II} = \begin{pmatrix} & & \text{Col}_i & & \text{Col}_j & & \\ & 1 & & & & & 0 \\ & & \ddots & & & & \\ & & & \boxed{0} & \cdots & \boxed{1} & \\ & & & \vdots & \ddots & \vdots & \\ & & & \boxed{1} & \cdots & \boxed{0} & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{pmatrix}$$

Since E_{II} is $I_{n \times n}$ with two rows interchanged, $\det(E_{II}) = -1$ and

$$\det(E_{II}A) = (-1) \cdot \det(A) = \det(E_{II}) \cdot \det(A)$$

- TYPE III. $(Row_i) \rightarrow (Row_i) + \lambda(Row_j)$, with $j \neq i$. Assuming $i < j$, the appropriate matrix is

$$E_{III} = \begin{pmatrix} & & \text{Col}_i & & \text{Col}_j & & \\ & 1 & & & & & 0 \\ & & \ddots & & & & \\ & & & \boxed{1} & \cdots & \boxed{\lambda} & \\ & & & \vdots & \ddots & \vdots & \\ & & & \boxed{0} & \cdots & \boxed{1} & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{pmatrix}$$

and we have $\det(E_{III}) = 1$. But then, we also have

$$\det(E_{III}A) = \det(A) = \det(E_{III}) \cdot \det(A)$$

This proves:

2.8. Lemma. *If E is any $(n \times n)$ elementary matrix then we have*

$$\det(EA) = \det(E) \cdot \det(A)$$

for any $n \times n$ matrix A .

This allows us to compute determinants using row operations, exploiting the fact that $\det(A)$ can be calculated by inspection if A is upper triangular. First observe that the effect of a sequence of row operations is to map $A \mapsto E_m \cdots E_1 \cdot A$ (echelon form), but then

$$\det(E_m \cdots E_1 A) = \det(E_m) \cdot \det(E_{m-1} \cdots E_1 \cdot A) = \left(\prod_{i=1}^m \det(E_i) \right) \cdot \det(A)$$

Thus

$$\det(A) = \left(\prod_{i=1}^m \det(E_i)^{-1} \right) \cdot \det(E_1 \cdots E_m A)$$

and calculating $\det(A)$ reduces to calculating the upper triangular row reduced form, whose determinant can be read by inspection. (You also have to keep track of the row

operations used, and their determinants.)

Computing Inverses. Suitably chosen row operations will put an $n \times n$ matrix into echelon form; if we only allow elementary operations of Type II or Type III we can achieve nearly the same result, except that the pivot entries contain nonzero scalars λ_i rather than “1”s, as shown in Figure 4.3. Next recall that $M(n, \mathbb{K})$ and the space of linear operators $\text{Hom}(\mathbb{K}^n, \mathbb{K}^n)$ are isomorphic as *associative algebras* under the correspondence

$$A \mapsto L_A \quad \left(L_A(\mathbf{x}) = A \cdot \mathbf{x} = ((n \times n) \cdot (n \times 1) \text{ matrix product}) \right),$$

as we showed in the discussion surrounding Exercise 4.12 of Chapter II. That means the following statements are equivalent.

- (32)
1. A matrix inverse A^{-1} exists in $M(n, \mathbb{K})$;
 2. $L_A : \mathbb{K}^n \rightarrow \mathbb{K}^n$ is an invertible linear operator;
 3. $\ker(L_A) = (0)$;
 4. The matrix equation $AX = 0$ has only the trivial solution $X = 0_{n \times 1}$.

We say that a matrix is **nonsingular** if any of these conditions holds; otherwise it is **singular**.

2.9. Exercise. If A, B are square matrices prove that

1. The product AB is *singular* if at least one of the factors is singular.
2. The product AB is *nonsingular* if both factors are nonsingular.

With this in mind we can deduce useful facts about matrix inverses from the preceding discussion of row operations and determinants.

2.10. Proposition. *The following statements regarding an $n \times n$ matrix are equivalent.*

1. $\det(A) \neq 0$;
2. A has a multiplicative inverse A^{-1} in $M(n, \mathbb{K})$;
3. The multiplication operator $L_A : \mathbb{K}^n \rightarrow \mathbb{K}^n$ is an invertible (bijective) linear operator on coordinate space.

Proof: We already know $(2.) \Leftrightarrow (3.)$. Row operations of Type II and III reduce A to one of the two “modified echelon forms” A' (see Figure 4.3(a–b)), in which the step corners contain nonzero scalars $\lambda_1, \dots, \lambda_r$ that need not equal 1, and $r = \text{rank}(A)$. Obviously if there are columns that do not meet a step-corner, as in 4.3(a), then the product of diagonal entries $\det(A)$ is zero; at the same time, the matrix equations $A'X = 0$ and $AX = 0$ will have nontrivial solutions, so the left multiplication operator $L_A : \mathbb{K}^n \rightarrow \mathbb{K}^n$ fails to be invertible (because $\ker(L_A) \neq (0)$) and a matrix inverse A^{-1} fails to exist. The situation in Figure 4.3(b) is better: since Type II and Type III operations can only change $\det(A)$ by a \pm sign, $\det(A) = \pm \det(A') = \pm \prod_{i=1}^n \lambda_i$ is nonzero. Concurrently, $AX = 0$ has only the trivial solution, L_A is an invertible linear operator on \mathbb{K}^n , and a matrix inverse A^{-1} exists. \square

To summarize, we have proved the following result (and a little more).

2.11. Theorem. *If $A \in M(n, \mathbb{K})$ then A^{-1} exists if and only if Type II and Type II row operations yield a modified echelon form that is upper triangular, with all diagonal*

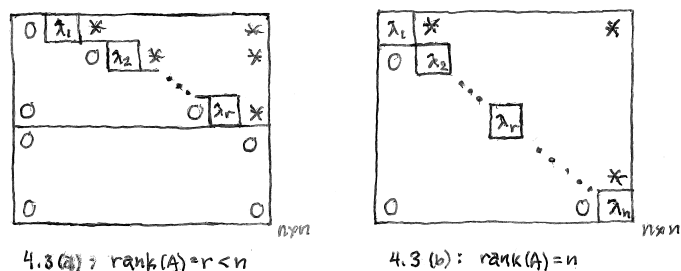


Figure 4.3. Row operations of Type II and III reduce an $n \times n$ matrix A to one of the two “modified echelon forms” A' shown in 4.3(a)–4.3(b); in both the step corners contain nonzero scalars $\lambda_1, \dots, \lambda_r$ that need not = 1, and $r = \text{rank}(A)$ with $r = n$ in 4.3(b).

If there are columns that do not meet a step-corner as in 4.3(a), then some diagonal entries in A' are zero and $\det(A) = \pm \det(A') = 0$. In the situation of 4.3(b) $\det(A) = \pm \det(A') = \pm(\lambda_1 \cdots \lambda_n)$ because Type II and III elementary operations have determinant = ± 1 . In this case $\det(A)$ is nonzero and its value can be determined by inspection, except for a (\pm) sign.

entries nonzero.

$$E_m \cdots E_1 A = \begin{pmatrix} \boxed{\lambda_1} & 0 & & * \\ 0 & \boxed{\lambda_2} & & \\ & \cdot & \cdot & \\ 0 & & 0 & \boxed{\lambda_n} \end{pmatrix} \quad (\lambda_i \neq 0)$$

Then the determinant is

$$\det(A) = \prod_{k=1}^m \det(E_k)^{-1} \cdot \prod_{i=1}^n \lambda_i$$

The factor $\prod_{k=0}^n \det(E_k)^{-1}$ attributed to the row operations can only be ± 1 since no Type I operations are involved. On the other hand, if the modified echelon form contains columns that do not meet a setp corner, then $\det(A) = 0$ and A^{-1} does not exist.

The basic definition (31) of the determinant is computationally very costly. Below we will give an algorithm (“*expansion by minors*”) which is often useful in studying the algebraic properties of determinants, but it is still pretty costly compared to the row reduction method developed above. To illustrate:

n = Matrix Size	Expansion by Minors		Row Reduction	
	Adds	Multiplies	Adds	Multiplies
2	1	2	1	3
4	23	40	14	23
5	119	205	30	45
10	3.6×10^6	6.2×10^6	285	339

The technique used above also yields a fairly efficient algorithm for computing A^{-1} (which at the same time determines whether A is in fact invertible). Allowing all three

types of row operations, an invertible matrix can be driven into its *reduced echelon form*, which is just the identity matrix $I_{n \times n}$. In this case

$$(33) \quad E_m \cdot \dots \cdot E_1 \cdot A = I_{n \times n} \quad \text{and} \quad A^{-1} = E_1^{-1} \cdot \dots \cdot E_m^{-1} \cdot I_{n \times n}$$

Each inverse E_k^{-1} is easily computed; it is just another elementary matrix of the same type as E_k . This can be codified as an explicit algorithm:

The Gauss-Seidel Algorithm. *Starting with the augmented $n \times 2n$ matrix $[A : I_{n \times n}]$, perform row operations to put A into “reduced” echelon form (upper triangular with zeros above all step corners). If $\text{rank}(A) < n$ and A is not invertible this will be evident – not all columns include a step-corner – and the algorithm reports that $\det(A) = 0$ and A is not invertible. Otherwise, every column is a pivot column and the reduced echelon form of A is just the identity matrix. Applying the same operations to the entire augmented matrix transforms $[A : I_{n \times n}] \rightarrow [I_{n \times n} : B]$ in which $B = A^{-1}$. (Why?)*

Another consequence of the preceding discussion is the very important multiplicative property of determinants.

2.12. Theorem (Multiplicative Property). *If $A, B \in M(n, \mathbb{K})$ then*

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

Proof: If A is singular then AB is singular (Exercise 2.9) and $\det(A) = 0$ (Exercise 2.10). Invoking Lemma 2.8 we get

$$\det(AB) = 0 = 0 \cdot \det(B) = \det(A) \cdot \det(B) \quad ,$$

and similarly if B is the singular factor.

Otherwise A and B are nonsingular and so is AB , so we can find elementary matrices such that $E_m \cdot \dots \cdot E_1 A = I_{n \times n}$, which implies $A = E_1^{-1} \cdot \dots \cdot E_m^{-1}$. By repeated application of Lemma 2.8 we see that

$$\det(A) = \prod_{i=1}^m \det(E_i^{-1})$$

and

$$\begin{aligned} \det(AB) &= \det(E_m^{-1}) \cdot \det(E_{m-1}^{-1} \cdot \dots \cdot E_1^{-1} B) \\ &= \dots \prod_i \det(E_i^{-1}) \cdot \det(B) = \det(A) \cdot \det(B). \quad \square \end{aligned}$$

2.13. Exercise. If $A \in M(n, \mathbb{K})$ is invertible then $\det(A^{-1}) = \det(A)^{-1}$. If $A, B \in M(n, \mathbb{K})$ and S is an invertible matrix such that $B = SAS^{-1}$ then $\det(B) = \det(A)$.

Thus $\det(A)$ is a “*similarity invariant*” – it has constant value for all matrices in a similarity class. We will encounter several other similarity invariants of matrices in the following discussion.

2.14. Exercise. Explain why $\text{rank}(A)$ of an $n \times n$ matrix is a similarity invariant.

2.15. Exercise. An $n \times n$ matrix A is said to be **orthogonal** if $A^t A = I_{n \times n}$. Prove that

1. $A^t A = I \Rightarrow AA^t = I$, so A is orthogonal $\Leftrightarrow A^t = A^{-1}$ (two-sided inverse).
2. $\det(A) = \pm 1$ for any orthogonal matrix, over any field.

Hint: Recall the comments posted in (32). For (1.) it suffices to show $A^t A = I \Rightarrow$ the operator $L_A : \mathbb{K}^n \rightarrow \mathbb{K}^n$ is one-to-one.

2.16. Exercise. Use Type II and III row operations to find the determinant of the following matrix.

$$A = \begin{pmatrix} 1 & 2 & 1 & 2 \\ 2 & 1 & 2 & 1 \\ 1 & 3 & 3 & 1 \\ 1 & 3 & 3 & 4 \end{pmatrix}$$

2.17. Exercise. Use Type II and III row operations to show that $\det(A) = -16i$ for the following matrix in $M(4, \mathbb{C})$, where $i = \sqrt{-1}$.

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}$$

2.18. Exercise. Apply the Gauss-Seidel algorithm to find A^{-1} for the matrices

$$(i) \ A = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 8 & 4 \\ 0 & 4 & 7 \end{pmatrix} \quad (ii) \ A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 4 & 1 \\ 0 & 4 & 2 \end{pmatrix}$$

2.19. Exercise. Consider the set of matrices H_n of the form

$$A = \begin{pmatrix} 1 & x_1 & \cdots & x_n & z \\ 0 & 1 & & & y_n \\ & & \ddots & & \vdots \\ & & & 1 & y_1 \\ 0 & & & & 1 \end{pmatrix}$$

with x_i, y_j, z in \mathbb{K} . When $\mathbb{K} = \mathbb{R}$ this is the n -dimensional *Heisenberg group* of quantum mechanics.

1. Prove that H_n is closed under matrix product.
2. Prove that the inverse A^{-1} of any matrix in H_n is also in H_n (compute it explicitly in terms of the parameters x_i, y_j, z).

Since the identity matrix is also in H_n , that means H_n is a *matrix group* contained in $GL(n+2, \mathbb{K})$.

2.20. Exercise. For $n \geq 2$ let

$$A = \begin{pmatrix} 0 & 1 & & & 0 \\ 1 & 0 & 1 & & \\ 0 & 1 & 0 & 1 & \\ & & & \ddots & \\ & & & 1 & 0 & 1 \\ 0 & & & & 1 & 0 \end{pmatrix}$$

Use row operations to

1. Calculate $\det(A)$.
2. Calculate the inverse A^{-1} if it exists.

Note: The outcome will depend on whether n is even or odd.

2.21. Exercise. Given a diagonal matrix $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ with distinct entries, find an invertible matrix S such that conjugation $D \mapsto SDS^{-1}$ interchanges the i^{th} and j^{th} diagonal entries ($i \neq j$):

$$S \begin{pmatrix} \lambda_1 & & & & 0 \\ & \ddots & & & \\ & & \boxed{\lambda_i} & & \\ & & & \ddots & \\ & & & & \boxed{\lambda_j} \\ & & & & & \ddots \\ 0 & & & & & & \lambda_n \end{pmatrix} S^{-1} = \begin{pmatrix} \lambda_1 & & & & 0 \\ & \ddots & & & \\ & & \boxed{\lambda_j} & & \\ & & & \ddots & \\ & & & & \boxed{\lambda_i} \\ & & & & & \ddots \\ 0 & & & & & & \lambda_n \end{pmatrix}$$

Hint: Think row *and* column operations on D . Note that if E_{II} is a Type II elementary matrix then $E^{-1} = E = E^t$, and right multiplication by E^t effects the corresponding column operation.

Determinants of Matrices vs Determinants of Linear Operators.

A determinant $\det(T)$ can be unambiguously assigned to any linear operator $T : V \rightarrow V$ on a finite dimensional space. Given a basis $\mathfrak{X} = \{e_i\}$ in V , we get a matrix $[T]_{\mathfrak{X}\mathfrak{X}}$ and could entertain the idea of assigning

$$(34) \quad \det(T) = \det([T]_{\mathfrak{X}\mathfrak{X}}) ,$$

but for this to make sense the outcome must be independent of the choice of basis. This actually works. If \mathfrak{Y} is any other basis we know there is an invertible matrix $S = [\text{id}_V]_{\mathfrak{Y}\mathfrak{X}}$ such that $[T]_{\mathfrak{Y}\mathfrak{Y}} = S [T]_{\mathfrak{X}\mathfrak{X}} S^{-1}$, and then by Theorem 2.12

$$\begin{aligned} \det([T]_{\mathfrak{Y}\mathfrak{Y}}) &= \det(S) \cdot \det([T]_{\mathfrak{X}\mathfrak{X}}) \cdot \det(S^{-1}) \\ &= \det(SS^{-1}) \cdot \det([T]_{\mathfrak{X}\mathfrak{X}}) = \det(I_{n \times n}) \cdot \det([T]_{\mathfrak{X}\mathfrak{X}}) \\ &= \det([T]_{\mathfrak{X}\mathfrak{X}}) \end{aligned}$$

as required. Thus the determinant (34) of a linear operator is well defined.

The **trace** $\text{Tr}(T)$ is another well-defined attribute of an operator $T : V \rightarrow V$ when $\dim(V) < \infty$. Recall Exercise 4.19 of Chapter II: For $n \times n$ matrices the trace $\text{Tr}(A) = \sum_{i=1}^n A_{ii}$ is a linear operator $\text{Tr} : M(n, \mathbb{K}) \rightarrow \mathbb{K}$ such that $\text{Tr}(I_{n \times n}) = n$ and $\text{Tr}(AB) = \text{Tr}(BA)$. If $\mathfrak{X}, \mathfrak{Y}$ are bases for V , we get

$$\text{Tr}([T]_{\mathfrak{Y}\mathfrak{Y}}) = \text{Tr}(S [T]_{\mathfrak{X}\mathfrak{X}} S^{-1}) = \text{Tr}(S^{-1} S \cdot [T]_{\mathfrak{X}\mathfrak{X}}) = \text{Tr}([T]_{\mathfrak{X}\mathfrak{X}})$$

Thus

$$(35) \quad \text{Tr}(T) = \text{Tr}([T]_{\mathfrak{X}\mathfrak{X}})$$

determines a well-defined trace on operators. Note, however, that if $T : V \rightarrow W$ with $V \neq W$, there is no natural way to assign a “determinant” or “trace” to T , even if $\dim(V) = \dim(W)$. The problem is philosophical: there is no natural way to say that a basis \mathfrak{X} in V is the “same as” another basis \mathfrak{Y} in W .

The operator trace has the same algebraic properties as the matrix trace.

2.22. Exercise. If $A, B : V \rightarrow V$ are linear operators on a finite dimensional space V , prove that

1. $\text{Tr} : \text{Hom}_{\mathbb{K}}(V, V) \rightarrow \mathbb{K}$ is a \mathbb{K} -linear map between vector spaces:

$$\text{Tr}(A + B) = \text{Tr}(A) + \text{Tr}(B) \quad \text{and} \quad \text{Tr}(\lambda \cdot A) = \lambda \cdot \text{Tr}(A)$$

2. $\text{Tr}(\text{id}_V) = n \cdot \dim(V)$;

3. $\text{Tr}(AB) = \text{Tr}(BA)$ (composition product of operators);

4. If S is an invertible operator and $B = SAS^{-1}$ then $\text{Tr}(B) = \text{Tr}(A)$.

The last statement shows that Tr is a similarity invariant for linear operators; so is the determinant \det .

2.23. Exercise. If $T : V \rightarrow V$ is a linear operator on a finite dimensional space prove that

$$\text{Tr}(T) = \text{Tr}(T^t) \quad \text{and} \quad \det(T) = \det(T^t)$$

Note: A conceptual issue arises here: T maps $V \rightarrow V$ while the transpose $T^t : V^* \rightarrow V^*$ acts on an entirely different vector space! But if you take a basis \mathfrak{X} in V and the dual basis \mathfrak{X}^* in V^* the definitions (34) and (35) still have something useful to say.

2.24. Exercise. Let $P : V \rightarrow V$ be a projection (associated with some direct sum decomposition $V = E \oplus F$) that projects vectors onto E along F . Prove that $\text{Tr}(P) = \dim_{\mathbb{K}}(E)$.

Hint: Pick a suitable basis compatible with the decomposition $E \oplus F$.

Expansion by Minors and Cramer's Rule. The following result allows a recursive computation of an $n \times n$ determinant once we can compute $(n-1) \times (n-1)$ determinants. Although it is useful for determining algebraic properties of determinants, and is handy for small matrices, it is prohibitively expensive in computing time for large n . This expansion is keyed to a particular row (or column) of A and involves an $(n-1) \times (n-1)$ determinant (the “minors” of the title) for each row entry.

2.25. Theorem (Cramer's Rule). For any row $1 \leq i \leq n$, we can write

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \cdot \det(\tilde{A}_{ij})$$

where \tilde{A}_{ij} = the $(n-1) \times (n-1)$ submatrix obtained by deleting Row_i and Col_j from A . Similarly, for any column $1 \leq j \leq n$ we have

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \cdot \det(\tilde{A}_{ij})$$

Proof: Since $\det(A) = \det(A^t)$, it is enough to prove the result for expansion along a row. Each term in the sum

$$\det(A) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \cdot (a_{1\sigma(1)} \cdots a_{n\sigma(n)})$$

contains just one term from $\text{Row}_i(A) = (a_{i1}, \dots, a_{in})$, so by gathering together terms we may write

$$\det(A) = a_{i1}a_{i1}^* + \dots + a_{in}a_{in}^*$$

in which a_{ij}^* involves no entry from $\text{Row}_i(A)$.

Our task is to show $a_{ij}^* = (-1)^{i+j} \det(\tilde{A}_{ij})$. One approach is to reduce to the case when $i = j = n$. In that special situation, we get

$$a_{nn}a_{nn}^* = \sum_{\sigma \in S'_n} \text{sgn}(\sigma) \cdot (a_{1\sigma(1)} \cdot \dots \cdot a_{n\sigma(n)})$$

where $S'_n \subseteq S_n$ is the subgroup of permutations such that $\sigma(n) = n$ (the subgroup that “stabilizes” the element “ n ” in $X = \{1, 2, \dots, n\}$).

2.26. Exercise. If $\tilde{\sigma} \in S_{n-1}$ is regarded as the permutation $\sigma \in S'_n \subseteq S_n$ such that $\sigma(n) = n$ and $\sigma(k) = \tilde{\sigma}(k)$ for $1 \leq k \leq n-1$, show that $\text{sgn}(\tilde{\sigma}) = \text{sgn}(\sigma)$.

In view of this the sum $\sum_{\sigma \in S'_n} (\dots)$ becomes $\sum_{\tilde{\sigma} \in S_{n-1}} (\dots)$. Thus

$$a_{nn}^* = (-1)^{n+n} \det(\tilde{A}_{nn}) = \det(A_{nn})$$

Now consider any i and j . Interchange $\text{Row}_i(A)$ with successive adjacent rows (“flips”) until it is at the bottom. This does not affect the value of $\det(\tilde{A}_{ij})$ because the relative positions of the other rows and columns are not affected; however each flip switched the sign of a_{ij} in the formula, and there are $n-i$ such changes. Similarly we may move $\text{Col}_j(A)$ to the n^{th} column, incurring $n-j$ sign changes. Thus

$$a_{ij}^* = (-1)^{n-i+n-j} \det(\tilde{A}_{ij}) = (-1)^{i+j} \det(\tilde{A}_{ij})$$

for all i and j , proving the theorem. \square

We post the following formula for A^{-1} without proof (cf Schaums, p 267-68). If matrix $A \in M(n, \mathbb{K})$ is invertible we have

$$(36) \quad A^{-1} = \frac{1}{\det(A)} \cdot (\text{Cof}(A))^t$$

where the $n \times n$ “*cofactor matrix*” $\text{Cof}(A)$ has i, j entry $= (-1)^{i+j} \tilde{A}_{ij}$, and \tilde{A}_{ij} = determinant of the $(n-1) \times (n-1)$ submatrix obtained by deleting (Row_i) and (Col_j) from A .