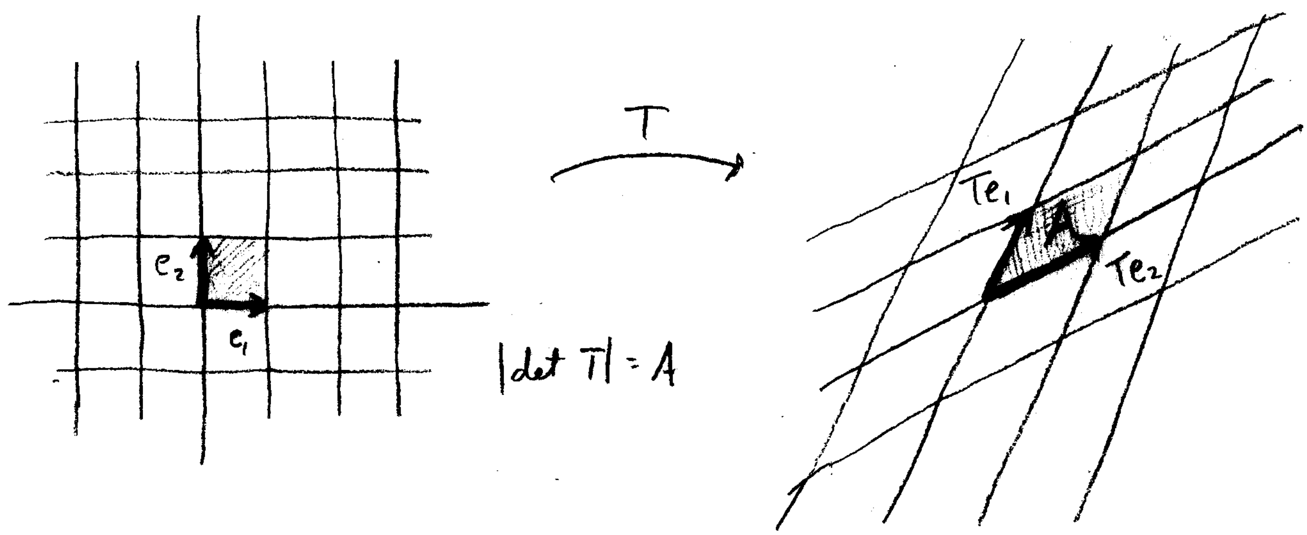
MTHSc 853: Linear Algebra

5. Determinant and trace.

The concept of the determinant of a linear map is simple – we will give an unofficial geometric definition as motivation, and then formalize it.

<u>Def</u>: (Unofficial). Let $T:\mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a linear map, and $[0,1]^n$ the unit n-cube. The determinant of T is the "signed volume" of $T([0,1]^n)$.



Our unofficial definition has some interesting properties:

- $\det T = 0$ iff T is not invertible
- $\det(TS) = (\det T)(\det S)$
- If T and S differ by swapping two columns, then $\det T = -\det S$.
- $\det T$ is "linear" in each column – i.e, if T is obtained from S by multiplying a column by c, then $\det T = c(\det S)$.

Permutations:

Def: Let $[n] := \{1, \dots, n\}$. A <u>permutation</u> is a bijection $\pi : [n] \longrightarrow [n]$. The set of all permutations of n elements is denoted S_n, and is a group.

We can describe a permutation by a table, or more concisely, by <u>cycle notation</u>:

<u>Example</u>: $\pi : \{1,2,3,4\} \longrightarrow \{1,2,3,4\}$.

i	1	2	3	4
$\pi(i)$	2	4	1	3

<u>Table notation</u>:

$$\pi = \frac{1234}{2413}, \quad \pi^2 = \frac{1234}{4321}, \quad \pi^3 = \frac{1234}{3142} = \pi^{-1}, \quad \pi^4 = \frac{1234}{1234}$$

<u>Cycle notation</u>: $\pi = (1\ 2\ 4\ 3)$ meaning $1 \to 2 \to 4 \to 3$

$$\pi^2 = (1\ 4)(2\ 3), \quad \pi^3 = (1\ 3\ 4\ 2), \quad \pi^4 = (1)(2)(3)(4).$$

By convention, we usually omit length-1 cycles, e.g. $(1 2)(3) = (1 2)$.

Def: Let $x_1, \dots, x_n$ be n variables. Their <u>discriminant</u> is defined as $P(x_1, \dots, x_n) = \prod_{i<j} (x_i - x_j)$.

For a permutation $\pi \in S_n$,

$$P(\pi(x_1, \dots, x_n)) := \prod_{i<j} (x_{\pi(i)} - x_{\pi(j)}) = \pm P(x_1, \dots, x_n).$$

Def: The <u>signature</u> $sgn(\pi)$ of a permutation $\pi \in S_n$ is defined as $P(\pi(x_1, \dots, x_n)) = sgn(\pi)\, P(x_1, \dots, x_n)$.

Clearly, $sgn(\pi) = \pm 1$.

Def: A <u>transposition</u> is a permutation $\tau \in S_n$ such that for some $j \neq k \in [n]$,

$$\tau(i) = i \quad i \neq j, k$$
$$\tau(j) = k$$
$$\tau(k) = j.$$

Prop: (i) $sgn(\pi_1 \circ \pi_2) = sgn(\pi_1)\, sgn(\pi_2)$

(ii) $sgn(\tau) = -1$ for any transposition

(iii) Every permutation $\pi \in S_n$ can be written as a composition of transpositions: $\pi = \tau_k \circ \dots \circ \tau_1$.

(iv) This decomposition is not unique, but the parity of k, the number of transpositions is.

(v) If $\pi = \tau_k \circ \dots \circ \tau_1$, then $sgn(\pi) = (-1)^k$.

Proof: Exercise.

<u>Multilinear forms</u>

Def: A <u>k-linear form</u> is a function $f: X_1 \oplus \dots \oplus X_k \longrightarrow K$ (we'll assume $X_1 = \dots = X_k = X$) that is linear in each coordinate, i.e, upon fixing any $k-1$ arguments, it remains linear in the remaining argument.

Examples:

(1) 1-linear forms are just functions

(2) 2-linear forms are bilinear forms

(3) A 3-linear form $f: X \oplus X \oplus X \longrightarrow K$ has identities such as $f(a_1x_1 + a_2x_2, y, z) = a_1 f(x_1, y, z) + a_2 f(x_2, y, z)$, and similarly for $f(x, a_1y_1 + a_2y_2, z)$, $f(x, y, a_1z_1 + a_2z_2)$, etc.

Theorem 5.1: The set of k-linear forms is a vector space of dimension n^k.

Proof: (sketch). Verify that a basis consists of functions

$$\{ F_{j_1 \cdots j_k}(X_{i_1}, \ldots, X_{i_k}) = \delta_{i_1 j_1} \cdots \delta_{i_k j_k} \; : \; 1 \le j_\ell \le n \}. \qquad \square$$

For any permutation $\pi \in S_k$, write

$$(\pi f)(x_1, \ldots, x_k) = f(x_{\pi(1)}, \ldots, x_{\pi(k)}).$$

Note that for any k-linear form f, πf is also k-linear.

Def: A k-linear form is symmetric if $\pi f = f$, for all $\pi \in S_k$.

Examples: (i) $f(x_1, x_2) = \ell_1(x_1)\ell_2(x_2) + \ell_1(x_2)\ell_2(x_1)$ for fixed $\ell_1, \ell_2 \in X'$.

(ii) $f(x_1, \ldots, x_k) = \sum_{\pi \in S_k} \pi f(x_1, \ldots, x_k)$

Def: A k-linear form is <u>skew-symmetric</u> if $\tau f = -f$ for every transposition $\tau \in S_k$.

Example: $f(x_1, x_2) = \ell_1(x_1)\ell_2(x_2) - \ell_1(x_2)\ell_2(x_1)$.

Def: A k-linear form is <u>alternating</u> if $f(x_1, \dots, x_k) = 0$ if $x_i = x_j$ for some $i \neq j$.

Prop: The set of alternating (resp. symmetric, or skew-symmetric) k-linear forms is a subspace.

Proof: Exercise.

Theorem 5.2: Every alternating multilinear form is skew-symmetric.

Proof: Pick $i \neq j$, and define $g(x_i, x_j) = f(x_1, \dots, x_k)$ (i.e, the other entries are fixed). Note that g is bilinear, alternating.

Thus,
$$0 = g(x_i + x_j, x_i + x_j) = g(x_i, x_i) + g(x_j, x_j) + g(x_i, x_j) + g(x_j, x_i)$$
$$= g(x_i, x_j) + g(x_j, x_i).$$

$\Rightarrow g(x_i, x_j) = -g(x_j, x_i)$

$\Rightarrow \tau f = -f$ for $\tau = (i\ j)$. $\square$

Remark: The converse "almost" holds.

Suppose $f = -f$. Then $(1+1)f = 0 \Rightarrow f = 0$ or $1+1=0$ (e.g., if $K = \mathbb{Z}_2$). Thus, the converse holds for $K \neq \mathbb{Z}_2$.

<u>Theorem 5.3</u>: If $x_1, \dots, x_k$ are linearly dependent and f is an alternating k-linear form, then $f(x_1, \dots, x_n) = 0$.

<u>Proof</u>: If $x_i = 0$, the result is trivial.

Otherwise, write $x_h = \sum_{i=1}^{h-1} a_i x_i$ for some x_h, and replace x_h in $f(x_1, \dots, x_k)$ by this sum.

Use linearity to break apart $f(x_1, \dots, x_k)$ into a sum, each one having some x_j in the h-position, $j \neq h$.

All of these terms are zero, since f is alternating. $\square$

The converse holds in one important special case: $k = n$.

<u>Theorem 5.4</u>: If f is a non-zero alternating n-linear form, and $x_1, \dots, x_n$ are linearly independent, then $f(x_1, \dots, x_n) \neq 0$.

<u>Proof</u>: By assumption, $x_1, \dots, x_n$ is a basis for X.

Pick $y_1, \dots, y_n \in X$, and write $y_i = a_{i_1} x_1 + \dots + a_{i_n} x_n$

Now,
$$f(y_1, \dots, y_n) = f\left(\sum a_{1_j} x_{1_j}, \dots, \sum a_{n_j} x_{n_j}\right)$$
$$= \sum_i f(z_{i_1}, \dots, z_{i_n}) \qquad z_{i_j} \text{ distinct } x_i\text{'s.}$$
$$= \sum_{\text{lots of } \pi} f(x_{\pi(1)}, \dots, x_{\pi(n)})$$
$$= \sum \operatorname{sgn}(\pi)\, f(x_1, \dots, x_n).$$

If $f(x_1, \dots, x_n) = 0 \Rightarrow f(y_1, \dots, y_n) = 0 \Rightarrow f = 0$ ↯. $\square$

<u>Remark</u>: The proof of Theorem 5.4 yields an important corollary:

<u>Theorem 5.5</u>: Any two alternating n-linear forms are linearly dependent.

<u>Proof</u>: Let f, g be alternating n-linear forms, and let $x_1, \dots, x_n$ be a basis of X.

Pick $y_1, \dots, y_n \in X$.

As before, $f(y_1, \dots, y_n) = \sum_{\text{lots}} \operatorname{sgn}(\pi)\, f(x_1, \dots, x_n) = a \in K$.

$$g(y_1, \dots, y_n) = \sum_{\text{lots}} \operatorname{sgn}(\pi)\, g(x_1, \dots, x_n) = b = \lambda a \in K.$$

Thus, $g(x_1, \dots, x_n) = \lambda f(x_1, \dots, x_n)$. $\square$

* Thus, the space of alternating n-linear forms is <u>at most</u>, one-dimensional.

<u>Theorem 5.6</u>: The vector space of alternating n-linear forms is one-dimensional.

<u>Proof</u>: It suffices to show that there exists one such form. We'll use induction on $k \le n$.

<u>Base case</u> ($k = 1$): Any $l \neq 0$ in X' is a non-trivial 1-linear form.

Now, suppose f is a non-zero alternating k-linear form, for $k < n$. We will construct a non-zero alternating $(k+1)$-linear form g.

8

Pick $y_1, \dots, y_k$ such that $f(y_1, \dots, y_k) \neq 0$, and find $y_{k+1} \in X$ not in the subspace $Y := \operatorname{span}\{y_1, \dots, y_k\}$.

Pick $\ell \in Y^\perp$ such that $\ell(y_{k+1}) \neq 0$, and $\ell(y_1) = \dots = \ell(y_k) = 0$.

Define: $g(x_1, \dots, x_k, x_{k+1}) = -f(x_1, \dots, x_k)\ell(x_{k+1}) + \sum_{i=1}^{k} (i\ \ k{+}1) f(x_1, \dots, x_k)\ell(x_{k+1})$

[For example, if $k=3$, then

$$g(x_1, x_2, x_3, x_4) = f(x_4, x_2, x_3)\,\ell(x_1) + f(x_1, x_4, x_3)\,\ell(x_2) + f(x_1, x_2, x_4)\,\ell(x_3) - f(x_1, x_2, x_3)\,\ell(x_4).]$$

It is elementary to prove that this is k-linear. We must show that it is non-zero and alternating.

Non-zero: $g(y_1, \dots, y_k, y_{k+1}) = -f(y_1, \dots, y_k)\ell(y_{k+1}) \neq 0$ ✓

Alternating: Consider vectors $x_1, \dots, x_k, x_{k+1}$ such that $i < j$ but $x_i = x_j$. It suffices to prove $g(x_1, \dots, x_k, x_{k+1}) = 0$.

Note: x_i and x_j occur in all but two "terms" of g. All other (besides these two) terms vanish.

Case 1: $(j = k+1)$:

$$g(x_1, \dots, x_{k+1}) = (i\ \ k{+}1) f(x_1, \dots, x_k)\ell(x_{k+1}) - f(x_1, \dots, x_k)\,\ell(x_{k+1})$$
$$= 0 \quad (\text{because } x_i = x_{k+1}).$$

Case 2: $(j \leq k)$:

$$g(x_1, \dots, x_{k+1}) = \left[(i\ \ j) f(x_1, \dots, x_k) - f(x_1, \dots, x_k)\right] \ell(x_{k+1})$$
$$= 0 \quad (\text{because } f \text{ is alternating}.$$

□

Let f be an alternating n-linear form on X, and $T: X \to X$ a linear map.

Note that $(\overline{T}f)(x_1, \dots, x_n) := f(Tx_1, \dots, Tx_n)$ is alternating and n-linear as well.

* Moreover, $\overline{T}$ is a linear map on the (one-dimensional) space of all such forms. Thus, $\overline{T}: f \longmapsto \lambda f$ for some $\lambda \in K$.

<u>Def</u>: This scalar is the <u>determinant</u> of T, denoted $\det T$.

Thus, the determinant satisfies the following:

<u>Universal property</u>: Given a linear map $T: X \longrightarrow X$, there exists a unique scalar $\lambda \in K$ such that for every alternating n-linear form f,

$$f(Tx_1, \dots, Tx_n) = \lambda f(x_1, \dots, x_n).$$

$$\begin{tikzcd} V^n \arrow[r, "T \oplus \cdots \oplus T"] \arrow[d, "f"'] & V^n \arrow[d, "f"] \\ K \arrow[r, "\lambda"] & K \end{tikzcd}$$

<u>Remark</u>: This definition is not only independent of matrices, but also independent of choice of basis!

<u>Example</u>: If $Tx = cx$, then

$$(\overline{T}f)(x_1, \dots, x_n) = f(cx_1, \dots, cx_n) = c^n f(x_1, \dots, x_n).$$

Thus, $\det T = c^n$. It follows that $\det 0 = 0$, $\det I = 1$.

<u>Theorem 5.7</u>: For any two linear maps $A, B : X \to X$,

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

<u>Proof</u>: Write $C = AB$. For an alternating n-linear form f;

$$(\bar{C}f)(x_1, \dots, x_n) = f(ABx_1, \dots, ABx_n)$$
$$= (\bar{A}f)(Bx_1, \dots, Bx_n) = (\bar{B}\bar{A}w)(x_1, \dots, x_n).$$

Thus, $\bar{C} = \bar{B}\bar{A}$. We have $\bar{C}f = (\det C)f$

and $\bar{C}f = \bar{B}\bar{A}f = (\det B)\bar{A}f = (\det B)(\det A)f$

$\Rightarrow \det(AB) = (\det A)(\det B)$. $\square$

<u>Corollary</u>: $A : X \to X$ is invertible iff $\det A \neq 0$.

<u>Proof</u>: If A is invertible, then $AA^{-1} = I$.

$1 = \det I = \det(AA^{-1}) = (\det A)(\det A^{-1})$. ✓

Conversely, if $\det A \neq 0$, and $x_1, \dots, x_n$ is a basis of X, and f a non-zero alternating n-linear form on X, then

$(\det A) f(x_1, \dots, x_n) \neq 0$ (by Theorem 5.4)

$\Rightarrow \{Ax_1, \dots, Ax_n\}$ is linearly independent (by Theorem 5.3)

$\Rightarrow A$ is invertible. $\square$

Determinants and matrices

Let f be a non-zero alternating n-linear form.

Let $\{x_1, \dots, x_n\}$ be a basis of X, and, let $A = (a_{ij})$ be the matrix of a linear map $X \to X$ wrt this basis.

Recall that $A x_j = \sum_{i=1}^{n} a_{ij} x_i$.

Thus, $(\det A) f(x_1, \dots, x_n) = f(A x_1, \dots, A x_n)$

$$= f\left(\sum_{i=1}^{n} a_{i1} x_i, \dots, \sum_{i=1}^{n} a_{in} x_i \right)$$

Now, expand this by multilinearity, to get a large sum of terms such as $f(z_1, \dots, z_n)$ where each z_i is a distinct x_i (if $z_i = z_j$ then $f(z_1, \dots, z_n) = 0$).

Moreover, every permutation $f(x_{\pi(1)}, \dots, x_{\pi(n)}) = \pi f(x_1, \dots, x_n)$ appears in this sum.

Remark: The coefficient of $\pi f(x_1, \dots, x_n)$ is $a_{\pi(1),1} \cdots a_{\pi(n),n}$.

Since f is skew-symmetric, we have

$$\det A = \sum_{\pi \in S_n} (\operatorname{sgn} \pi)\, a_{\pi(1),1} \cdots a_{\pi(n),n}. \qquad (*)$$

Also note that if $\pi, \sigma \in S_n$, then $a_{\pi(1),1} \cdots a_{\pi(n),n}$ and $a_{\pi\sigma(1),\sigma(1)} \cdots a_{\pi\sigma(n),\sigma(n)}$ differ in the order of the factors only.

Let's apply this with $\sigma = \pi^{-1}$. (Note: $\mathrm{sgn}(\pi) = \mathrm{sgn}(\pi^{-1})$)

$$\det A = \sum_{\pi \in S_n} (\mathrm{sgn}\, \pi)\, a_{\pi(1),1} \cdots a_{\pi(n),n} = \sum_{\pi \in S_n} (\mathrm{sgn}\, \pi)\, a_{1,\pi(1)} \cdots a_{n,\pi(n)}.$$
$$= \det A^T.$$

* In summary, the determinant of A can be thought of as an n-linear function of its column vectors, or of its row vectors. In fact it is the unique "normalized" alternating n-linear form, so that $f(e_1, \ldots, e_n) = 1$. ($\det I = 1$).

<u>Lemma 5.8</u>: Let $A = (c_1, \ldots, c_n)$ be a matrix (c_i is a column vector) and let B be the matrix obtained from A by adding k times the i^{th} column of A to the j^{th} column, $i \neq j$. Then $\det A = \det B$.

<u>Proof</u>: Exercise.

<u>Lemma 5.9</u>: Let A be an $n \times n$ matrix whose first column is e_1:

$$A = \begin{pmatrix} 1 & x & x & x \\ 0 & & & \\ \vdots & & A_{11} & \\ 0 & & & \end{pmatrix}.$$

Then $\det A = \det A_{11}$

<u>Proof</u>: By lemma 5.8, $\det A = \det \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & A_{11} \\ 0 & \end{pmatrix}$.

Define $f(A_{11}) = \det \begin{pmatrix} 1 & 0 \\ 0 & A_{11} \end{pmatrix}$, a function of an $(n-1)\times(n-1)$ matrix A_{11}.

Clearly, f is an alternating n-linear form with $f(I)=1$, so it must be the determinant function. □

<u>Corollary 5.10</u>: Let A be a matrix whose j^{th} column is e_i. Then $\det A = (-1)^{i+j} \det A_{ij}$, where A_{ij} is the $(n-1)\times(n-1)$ matrix obtained by striking out the i^{th} row and j^{th} column of A (called the $(ij)^{th}$ minor of A).

<u>Proof</u>: Exercise.

<u>Theorem 5.11</u>: Let A be an $n\times n$ matrix, and $1 \le j \le n$. Then $\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} (\det A_{ij})$. (The "<u>Laplace expansion</u>.")

<u>Proof</u>: To simplify notation, take $j=1$.

Write $a_1 = a_{11}e_1 + \cdots + a_{n1}e_n$, where $A = (a_1, \ldots, a_n)$.

By multilinearity, $\det A = f(a_1, \ldots, a_n)$

$$= f(a_{11}e_1 + \cdots + a_{n1}e_n,\ a_2, \ldots, a_n)$$

$$= a_{11} f(e_1, a_2, \ldots, a_n) + \cdots + a_{n1} f(e_n, a_2, \ldots, a_n)$$

Now apply Corollary 5.10. □

<u>Application</u>: Systems of equations.

Consider a system $Ax = u$, A is invertible and $x = \sum_{j=1}^{n} x_j e_j$,

$$A = (a_1, \ldots, a_n) = (Ae_1, \ldots, Ae_n).$$

Thus, our system can be written as $\sum_{j=1}^{n} x_j a_j = u$.

Let $A_k = (a_1, \dots, a_{k-1}, u, a_{k+1}, \dots, a_n)$ for each k.

$$= \left(a_1, \dots, a_{k-1}, \sum_{j=1}^{n} x_j a_j, a_{k+1}, \dots, a_n\right).$$

By multilinearity, $\det A_k = \sum_{j=1}^{n} x_j \det(a_1, \dots, a_{k-1}, a_j, a_{k+1}, \dots, a_n)$

$$= x_k \det(a_1, \dots, a_{k-1}, a_k, a_{k+1}, \dots, a_n)$$

$$= x_k \det A.$$

Since A is invertible, $\det A \neq 0 \implies x_k = \dfrac{\det A_k}{\det A}$.

Now apply the Laplace Expansion (Thm 5.11) to A_k:

$$\det A_k = \sum_{i=1}^{n} (-1)^{i+k} \det A_{ik}\, u_i$$

$\implies x_k = \sum_{i=1}^{n} (-1)^{i+k} \dfrac{\det A_{ik}}{\det A} u_i$. This is "<u>Cramer's rule</u>."

<u>Theorem 5.12</u>: If A is invertible, then its inverse A^{-1} has the form $(A^{-1})_{ki} = (-1)^{i+k} \dfrac{\det A_{ik}}{\det A}$.

<u>Proof</u>: Consider a system $Ax = u \implies x = A^{-1}u$. By Cramer's rule:

$$\sum_{i=1}^{n} (-1)^{i+k} \frac{\det A_{ik}}{\det A} u_i = x_k = (A^{-1}u)_k = \sum_{i=1}^{n} (A^{-1})_{ki} u_i.$$

$\square$

<u>Note</u>: This formula isn't "practical" for computing inverses.

<u>Def</u>: The <u>trace</u> of an $n \times n$ matrix is $\mathrm{tr}\, A = \sum_{i=1}^{n} a_{ii}$.

<u>Theorem 5.13</u>:

(a) Trace is linear: $\mathrm{tr}(kA) = k(\mathrm{tr}\, A)$

$\mathrm{tr}(A+B) = \mathrm{tr}\, A + \mathrm{tr}\, B$.

(b) Trace is "commutative": $\mathrm{tr}(AB) = \mathrm{tr}(BA)$

<u>Proof</u>: (a) is obvious.

For (b): $(AB)_{ii} = \sum_{k=1}^{n} a_{ik} b_{ki}$ and $(BA)_{ii} = \sum_{k=1}^{n} b_{ik} a_{ki}$

Thus, $\mathrm{tr}(AB) = \sum_{i,k} a_{ik} b_{ki} = \sum_{i,k} b_{ik} a_{ki} = \mathrm{tr}(BA)$. □

<u>Theorem 5.14</u>: Similar matrices have the same determinant and trace.

<u>Proof</u>: Suppose that $A = SBS^{-1}$.

$\det A = \det(SBS^{-1}) = (\det S)(\det B)(\det S^{-1}) = (\det B)(\det S)(\det S^{-1})$

$= (\det B) \det(SS^{-1}) = \det B$. ✓

$\mathrm{tr}\, A = \mathrm{tr}(SBS^{-1}) = \mathrm{tr}(S^{-1}SB) = \mathrm{tr}\, B$. ✓

□

Since similar matrices represent the same linear map but with a different choice of basis, it is well-founded to speak of the determinant and trace as functions of linear maps, not just matrices.