Determinants Lecture

Sign of a Permutation

Let $\sigma \in S_n$ be given. Let $c(\sigma)$ denote the number of disjoint cycles of σ. We define the sign of σ by

$$\text{sgn}(\sigma) = (-1)^{n-c(\sigma)}.$$

In particular, if σ is a 2-cycle, then

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

Theorem:

(a) Let $\tau_1, \ldots, \tau_k$ be 2-cycles. Then $\text{sgn}(\tau_1 \cdots \tau_k) = (-1)^k$.

(b) Let σ, τ be permutations. Then $\text{sgn}(\sigma\tau) = \text{sgn}(\sigma)\text{sgn}(\tau)$.

Proof: Observe that we have

$$(axby)(ab) = (ay)(bx)$$

and

$$(ax)(by)(ab) = (aybx)$$

where a and b are distinct elements of $\{1, 2, \ldots, n\}$ and x represents a sequence of elements from $\{1, 2, \ldots, n\} \backslash \{a, b\}$. These two formulas implies that if σ has c disjoint cycles and $\tau = (ab)$ then $\sigma\tau$ has either $c+1$ disjoint cycles or $c - 1$ disjoint cycles, depending on whether a and b appear in the same cycle of σ or two different cycles of σ. Hence

$$\text{sgn}(\sigma\tau) = -\text{sgn}(\sigma) = \text{sgn}(\sigma)\text{sgn}(\tau)$$

when τ is a 2-cycle. This implies that

$$\text{sgn}(\tau_1 \cdots \tau_k) = (-1)^k,$$

when $\tau_1, \ldots, \tau_k$ are 2-cycles. Since every permutation can be factored into 2-cycles, this yields (a), and (b) follows from (a).

Determinant of a Matrix

$$\det A = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) A(\sigma),$$

where $\mathcal{S}_n$ is the set of permutations of $(1, 2, \ldots, n)$ and

$$A(\sigma) = a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

In particular, $\det I = 1$.

Theorem: $\det A^T = \det A$.

Proof:

$$\det A^T = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) A^T(\sigma) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma^{-1}) A(\sigma^{-1}) =$$

$$\sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) A(\sigma) = \det A.$$

Matrices can be regarded as lists of columns: $A = (A_1, A_2, \ldots, A_n)$. The next theorem says that, as a function of column lists, the determinant is multilinear and skew-symmetric:

Theorem:

(1) $\det(*, A+B, *) = \det(*, A, *) + \det(*, B, *)$ and $\det(*, \lambda A, *) = \lambda \det(*, A, *)$.

(2) If $A_i = A_j$ for some $i \neq j$ then $\det(A_1, \ldots, A_n) = 0$.

(3) $\det(*, A, *, B, *) = -\det\,(*, B, *, A, *)$.

Proof: Property (1) is proved by a direct calculation.

To prove (2), let S be the set of all permutations with $\sigma(i) > \sigma(j)$ and let T be the set of all permutations with $\tau(i) < \tau(j)$. There is a one-to-one correspondence between S and T via $\sigma \mapsto \sigma(i, j)$. When $A_i = A_j$ we have $A(\sigma) = A(\tau)$. Hence the terms in the determinant expansion can be grouped into pairs with opposite signs and they all cancel out.

To prove (3), we use (1) and (2):

$$\det(*, A, *, B, *) + \det(*, B, *, A, *) =$$

$$\det(*, A, *, A, *) + \det(*, A, *, B, *) + \det(*, B, *, A, *,) + \det(*, B, *, B, *) =$$

$$\det(*, A+B, *, A+B, *) = 0.$$

Corollary: If the columns $(A_1, \ldots, A_n)$ are linearly dependent then

$$\det(A_1, \ldots, A_n) = 0.$$

Proof: Let's say that column A_i is a linear combination of the other columns:

$$A_i = \sum_{p \neq i} \alpha_p A_p.$$

By multilinearity we have

$$\det(A_1, \ldots, A_n) = \sum_{p \neq i} \alpha_p \det(A_1, \ldots, A_p, \ldots, A_n).$$

Since the terms determinants in the sum operate on lists with a repeated column, the sum is zero.

Theorem: For any pair of $n \times n$ matrices A and B, $\det(AB) = \det(A)\det(B)$.

Proof: Let $A = (a_{ij})$ and $B = (b_{ij})$ be given. Then AB has columns $(C_1, \ldots, C_n)$, where

$$C_i = b_{1i}A_1 + b_{2i}A_2 \cdots + b_{ni}A_n.$$

Using the multilinearity of the determinant, we have

$$\det(AB) = \det(\sum_i b_{1i}A_i, \sum_i b_{2i}A_i, \ldots, \sum_i b_{ni}A_i) =$$

$$\sum_{i_1, i_2, \ldots, i_n} b_{1i_1} b_{2i_2} \cdots b_{ni_n} \det(A_{i_1}, A_{i_2}, \ldots, A_{i_n}).$$

Instances of repeated columns among the $(A_{i_1}, A_{i_2}, \ldots, A_{i_n})$ contribute zero to the sum, so we can assume that we are only dealing with the lists $(A_{\sigma(1)}, \ldots, A_{\sigma(n)})$ for permutations $\sigma \in \mathcal{S}_n$, in which case

$$\det(A_{\sigma(1)}, \ldots, A_{\sigma(n)}) = \mathrm{sgn}(\sigma)\det(A_1, \ldots, A_n) = \mathrm{sgn}(\sigma)\mathrm{det}\ A.$$

Hence

$$\det(AB) = \det A \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) b_{1\sigma(1)} \cdots b_{n\sigma(n)} = \det(A)\det(B).$$

Theorem: A matrix A is invertible if and only if $\det A \neq 0$.

Proof: If A is not invertible then then its columns are linearly dependent, and the corollary yields $\det(A_1, \ldots, A_n) = 0$. If A is invertible, then $AB = I$ is possible. Therefore $\det(A)\det(B) = \det(I) = 1$, which implies that $\det(A) \neq 0$.

Cramer's Rule: Consider the matrix equation $Ax = b$ where A is a square invertible matrix. The unique solution to this equation is $x = A^{-1}b$. Let the coordinates of x be $x_1, x_2, \ldots, x_n$. We can express each of these numbers in terms of determinants as follow:

Let the columns of A be $A_1, \ldots, A_n$. Then $b = x_1 A_1 + \cdots + x_n A_n$. Consider the matrix $A^{(i)}$ which results after replacing column A_i by b. Then

$$\det A^{(i)} = \det(A_1, \ldots, b, \ldots, A_n) = \det(A_1, \ldots, x_1 A_1 + \cdots + x_n A_n, \ldots, A_n) =$$

$$x_1 \det(A_1, \ldots, A_1, \ldots, A_n) + x_2 \det(A_1, \ldots, A_2, \ldots, A_n) + \cdots$$
$$+ x_n \det(A_1, \ldots, A_n, \ldots, A_n).$$

The only surviving terms is

$$x_i \det(A_1, \ldots, A_i, \ldots, A_n) = x_i \det A.$$

Therefore

$$x_i = \frac{\det A^{(i)}}{\det A}.$$

Row-Expansion of a Determinant

Let A be a matrix. Let A_{ij} denote the matrix obtained by deleting row i and column j. For any p,

$$\det A = \sum_{q=1}^{n} (-1)^{p-q} a_{pq} \det(A_{pq}).$$

Proof: The determinant of A is

$$\sum_{\sigma} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \sum_{q=1}^{n} \sum_{\substack{\sigma \\ \sigma(p)=q}} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}.$$

Let τ be the permutation that exchanges p and q and leaves the other indices fixed. Set $b_{ij} = a_{i\tau(j)}$ for all i and j. Then we have $a_{i\sigma(i)} = b_{i\tau\sigma(i)}$ and we can write

$$\sum_{\substack{\sigma \\ \sigma(p)=q}} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \sum_{\substack{\sigma \\ \sigma(p)=q}} \operatorname{sgn}(\sigma) b_{1\tau\sigma(1)} \cdots b_{n\tau\sigma(n)} =$$

$$\operatorname{sgn}(\tau) b_{pp} \sum_{\substack{\sigma \\ \sigma(p)=q}} \operatorname{sgn}(\tau\sigma) b_{1\tau\sigma(1)} \cdots \widehat{b_{pp}} \cdots b_{n\tau\sigma(n)} =$$

$$\operatorname{sgn}(\tau) a_{pq} \sum_{\substack{\gamma \\ \gamma(p)=p}} \operatorname{sgn}(\gamma) b_{1\gamma(1)} \cdots \widehat{b_{pp}} \cdots b_{n\gamma(n)} =$$

$$\operatorname{sgn}(\tau) a_{pq} \det(B_{[n]\setminus\{p\}}).$$

If $p = q$ then $B_{[n]\setminus\{p\}} = A_{pp}$, and if $p \neq q$ then $B_{[n]\setminus\{p\}}$ can be obtained from A_{pq} by making a series of $|p - q| - 1$ swaps in its columns. In either case we have

$$\operatorname{sgn}(\tau) a_{pq} \det(B_{[n]\setminus\{p\}}) = (-1)^{p-q} a_{pq} \det(A_{pq}).$$

Theorem: Let A be an invertible matrix. Let

$$b_{ij} = \frac{(-1)^{i+j} \det\ A_{ji}}{\det\ A}$$

for each i and j. Then $A^{-1} = (b_{ij})$.

Proof: The pq element of AB is

$$\sum_{i} a_{pi} b_{iq} = \frac{1}{\det\ A} \sum_{i} a_{pi} (-1)^{i+q} \det\ A_{qi}.$$

This is equal to $\frac{1}{\det\ A}$ times the row-q expansion of the determinant of the matrix A' obtained from A by replacing row q by row p. If $p \neq q$ then A'

has a repeated row, hence det $A' = 0$. But when $p = q$ we have $A' = A$ and det $A' =$ det A. Therefore the pq element of AB is 0 if it is not on the main diagonal, 1 if it is on the main diagonal. Hence $AB = I$.

Change of Basis Matrix

Let $S, T \in \mathcal{L}(v)$ be given. Suppose that $A = M(T)$ with respect to the basis $(v_1, \ldots, v_n)$ and that $B = M(T)$ with respect to the basis $(w_1, \ldots, w_n)$. How are the matrices A and B related? Let $R : V \to V$ be the linear operator defined by $R(v_i) = w_i$ for $i \leq n$. Let $C = M(R)$ with respect to the basis $(v_1, \ldots, v_n)$. Since R is invertible, so is C. We have

$$Tw_i = \sum_p b_{pi} w_p.$$

Therefore

$$TRv_i = \sum_p b_{pi} w_p.$$

Therefore

$$R^{-1}TRv_i = \sum_p b_{pi} v_p.$$

Therefore B represents $R^{-1}TR$ with respect to the basis $(v_1, \ldots, v_n)$. On the other hand, so does $C^{-1}AC$. Therefore we have

$$C^{-1}AC = B.$$

A more suggestive notation is

$$M_v(w, v)M(T, v)M_v(v, w) = M(T, w)$$

where $M(T, v) = A$, $M(T, w) = B$, $M_v(v, w) = C$, $M_v(w, v) = C^{-1}$.

Determinant of an endomorphism: Let V be a finite-dimensional vector space and let $f : V \to V$ be an endomorphism. We proved above that if A is a matrix representing f with respect to one basis and B is a matrix representing f with respect to a second basis, then $C^{-1}AC = B$ for some matrix invertible matrix C. This implies that det $A =$ det B. Hence there is an unambiguous number we can attach to f which we can call det f, namely the common value of all its matrix representations.

Determinant of $T \in \mathcal{L}(F^n)$ where $F \in \mathbb{R}, \mathbb{C}$: Let $T = S\sqrt{T^*T}$ be the singular-value decomposition of T. Since $S^*S = I$, $|\det(S)| = 1$. Therefore $|\det(T)| = |\det(\sqrt{T^*T})| = \det(\sqrt{T^*T})$.

Trace of a square matrix: Let $A = (a_{ij})$ be an $n \times n$ matrix. Then the trace of A is

$$\text{trace}(A) = a_{11} + a_{22} + \cdots + a_{nn}.$$

Trace of $T \in \mathcal{L}(V)$: It is easy to verify that $\text{trace}(AB) = \text{trace}(BA)$ for any pair of $n \times n$ matrices A and B. In particular, for any invertible matrix C, $\text{trace}(C^{-1}AC) = \text{trace}(AC^{-1}C) = \text{trace}(A)$. Hence every matrix representation of T with respect a basis for V has the same trace, and we call this the trace of T.

Volume: A product of intervals $\Omega = [a_1, b_1] \times \cdots \times [a_n, b_n] \subseteq \mathbb{R}^n$ is called a rectangular solid. We define the volume of Ω to be $\text{vol}(\Omega) = \prod_{i=1}^{n} (b_i - a_i)$. We will also define

$$\text{vol}(\Omega_1 \cup \cdots \Omega_k) = \text{vol}(\Omega_1) + \cdots + \text{vol}(\Omega_k)$$

whenever $\Omega_1 \cup \cdots \Omega_k$ is a disjoint union. Given a subset X of $\mathbb{R}^n$, we will define $v(X)$ to be the supremum of all $\text{vol}(\Omega_1 \cup \cdots \Omega_k)$ where $\Omega_1 \cup \cdots \Omega_k \subseteq X$ is a disjoint union of rectangular solids in X, and we will define $V(X)$ to be the infimum of all $\text{vol}(\Omega_1 \cup \cdots \Omega_k)$ where $\Omega_1 \cup \cdots \Omega_k \supseteq$ is disjoint union of rectangular solids containing X. If $v(X) = V(X) = v$ then v must be a real number we define $\text{vol}(X) = v$.

Computing vol$(T(X))$ where $T \in \mathcal{L}(V)$ and $V \in \{\mathbb{R}^2, \mathbb{R}^3\}$:

1. If T is a positive linear operator with diagonal matrix representation

$$\text{diag}(r_{11}, \ldots, r_{nn})$$

with respect to the standard basis $\{e_1, \ldots, e_n\}$ then

$$\text{vol}(T(\Omega)) = r_{11} \cdots r_{nn} \text{vol}(\Omega) = \det(T)\text{vol}(\Omega)$$

since all T does is scale up the dimensions of Ω by its diagonal entries. This implies that

$$\text{vol}(T(X)) = \det(T)\text{vol}(X)$$

using the definition of $\text{vol}(X)$ and properties of the infimum and supremum.

2. If $S : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry then, since it preserves lengths of vectors and angles between vectors,

$$\text{vol}(S(\Omega)) = \text{vol}(\Omega).$$

This implies

$$\text{vol}(S(X)) = \text{vol}(X)$$

for any X whose volume can be measured, using the definition of $v(X)$ and properties of the infimum and supremum.

3. If T is a positive linear operator then

$$\text{vol}(T(X)) = \det(T)\text{vol}(X).$$

Proof: Say that T has diagonal matrix representation

$$\text{diag}(r_{11}, \ldots, r_{nn})$$

with respect to an orthonormal basis $\{f_1, \ldots, f_n\}$. Let S be the isometry that maps e_i to f_i for each i, where $e_1, \ldots, e_n$ is the standard basis. Then $S^{-1}TS$ has the same matrix representation with respect to the standard basis. It is easy to verify that $S^{-1}TS$ is a positive linear operator. By #1 above,

$$\text{vol}(S^{-1}TS(X)) = \det(S^{-1}TS)\text{vol}(X) = \det(T)\text{vol}(X).$$

On the other hand, by #2 above we also have

$$\text{vol}(S^{-1}TS(X)) = \text{vol}(TS(X)).$$

Taken together,

$$\text{vol}(TS(X)) = \det(T)\text{vol}(X)$$

for any X. Therefore

$$\text{vol}(T(X)) = \det(T)\text{vol}(S^{-1}(X)) = \det(T)\text{vol}(X).$$

4. If T is an arbitrary linear operator then

$$\text{vol}(T(X)) = |\det(T)|\text{vol}(X).$$

Proof: Write $T = S\sqrt{T^*T}$ where S is an isometry. Then

$$\mathrm{vol}(T(X)) = \mathrm{vol}(\sqrt{T^*T}(X)) = \det(\sqrt{T^*T})\mathrm{vol}(X) = |\det(T)|\mathrm{vol}(X).$$

Example: Area of the ellipse $E = \{(x,y) : \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1\}$**.**

The ellipse is the image of the unit disk $D = \{(x,y) : x^2 + y^2 = 1\}$ under the mapping $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x,y) = (ax, by)$ where $a, b > 0$. Therefore

$$\mathrm{vol}(E) = \mathrm{vol}(T(C)) = \det(T)\mathrm{vol}(D) = ab\pi.$$

Change of variables formula: The volume integral of a continuous function $f : \mathbb{R}^3 \to \mathbb{R}$ over a region $X \subseteq \mathbb{R}^3$ can be expressed as the limit in some sense of all expressions of the form

$$\sum f(a_i, b_i, c_i)\mathrm{vol}(X_i)$$

where $X_1 \cup \cdots X_k \subseteq X$ is a disjoint union of small regions and $(a_i, b_i, c_i) \in X_i$ for each i. Now suppose there is a differentiable injective mapping from Y onto X via ϕ. We can express the integral as the limit of expressions of the form

$$\sum f(\phi(a_i, b_i, c_i))\mathrm{vol}(\phi(\Omega_i))$$

where $\Omega_1 \cup \cdots \Omega_k \subseteq Y$ is a disjoint union of small regions and $(a_i, b_i, c_i) \in \Omega_i$ for each i. For Ω with small dimensions, the region $\phi(\Omega)$ is approximated by the region $\left(\frac{\partial \phi_i(a_i,b_i,c_i)}{\partial x_j}\right)(\Omega)$, evaluated at some $(a_i, b_i, c_i) \in \Omega_i$, yielding the approximation

$$\sum f(\phi(a_i, b_i, c_i)) \left|\det\left(\frac{\partial \phi_i(a_i, b_i, c_i)}{\partial x_j}\right)\right| \mathrm{vol}(\Omega_i).$$

Hence we obtain

$$\int\int\int_X f(x,y,z)\, dV = \int\int\int_Y f(\phi(x,y,z)) \left|\det\left(\frac{\partial \phi_i}{\partial x_j}\right)\right|\, dV.$$

If Y has has a simple shape then the latter can be evaluated by an iterated integral.

For example, if X is the upper-half of the unit sphere and $Y = [0,1] \times [0, 2\pi] \times [0, \frac{\pi}{2}]$ then $\phi : Y \to X$ defined by $\phi(x,y,z) = (x \sin z \cos y, x \sin z \sin y, x \cos z)$ maps Y onto X. We have