ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

IFOR
Institute for Operations Research
Department of Mathematics

Robert Weismantel

Lecture 22: Determinants

The determinant as a function over matrices

A function defined for square matrices

Let $A \in \mathbb{R}^{n \times n}$ be a matrix. We consider the function $\det(A) \in \mathbb{R}$. The function value $|det(A)|$ measures the volume of the parallelopiped

$$\mathscr{P} = \text{vol}(\{x \in \mathbb{R}^n \mid \exists 0 \leq \lambda_i \leq 1 \text{ for } i = 1, \ldots, n \text{ such that } x = \sum_{i=1}^{n} \lambda_i A_{\cdot,i}\}).$$

The following properties hold.

- $A \in \mathbb{R}^{n \times n}$ is invertible if and only if $\det(A) \neq 0$.
- $\det(A) = \det(A^T)$.
- Linearity: Let A and B be two matrices in $\mathbb{R}^{n \times n}$ where all rows are equal except for row i. Let C be the matrix with rows $C_{j,\cdot} = A_{j,\cdot} = B_{j,\cdot}$ for all $j \in \{1, \ldots, n\} \setminus \{i\}$ and $C_{i,\cdot} = A_{i,\cdot} + B_{i,\cdot}$. Then

$$\det(C) = \det(A) + \det(B).$$

- For matrices A and B in $\mathbb{R}^{n \times n}$ $\det(AB) = \det(A)\det(B)$.

2×2- matrices

The determinant

- For $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$ we define $det(A) = ad - bc$.
- For 2-by-2 matrices A, W we have $\det(AW) = \det(A)\det(W)$.

Proof.

- $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \quad W = \begin{bmatrix} x & z \\ y & w \end{bmatrix} \quad AW = \begin{bmatrix} ax+cy & az+cw \\ bx+dy & bz+dw \end{bmatrix}.$
- $$\begin{aligned} det(AW) &= (ax+cy)(bz+dw)-(az+cw)(bx+dy) \\ &= axbz+axdw+cybz+cydw \\ &\quad -azbx-azdy-cwbx-cwdy \end{aligned}$$
- $$\begin{aligned} det(AW) &= axdw+cybz-azdy-cwbx \\ &= ad(xw-zy)+cb(zy-xw) \\ &= \det(A)\det(W). \end{aligned}$$

□

This computation characterizes when a 2×2-matrix is invertible.

Lemma

A matrix $A \in \mathbb{R}^{2\times 2}$ is invertible if and only if $\det(A) \neq 0$.

Proof.

$$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}.$$

- If A is invertible, then $AA^{-1} = I$ implies $\det(A)\det(A^{-1}) = 1$. Hence, $\det(A) \neq 0$.
- For the converse direction, assume $\det(A) \neq 0$, i.e., $a \neq 0$ or $b \neq 0$. Wlog. $a \neq 0$. Consider the system of linear equations $AW = I$.

$$\begin{array}{lll} ax + cy = 1 & \text{implies that} & x = \frac{1-cy}{a} \\ az + cw = 0 & \text{implies that} & z = \frac{-cw}{a}. \end{array}$$

Proof continued

- By substituting $x = \frac{1-cy}{a}$ and $z = \frac{-cw}{a}$ into $bx + dy = 0$ we get

$$\frac{b}{a} - \frac{cyb}{a} + dy = 0 \iff b + y(ad - bc) = 0 \iff y = \frac{-b}{\det(A)}.$$

- By substituting $x = \frac{1-cy}{a}$ and $z = \frac{-cw}{a}$ into $bz + dw = 1$ we obtain

$$\frac{-bcw}{a} + dw = 1 \iff -bcw + adw = a \iff w = \frac{a}{\det(A)}.$$

- This gives us a formula for the parameters z and x in form of

$$z = \frac{-c}{\det(A)} \text{ and } x = \frac{1 + \frac{cb}{\det(A)}}{a} = \frac{ad - bc + cb}{a\det(A)} = \frac{d}{\det(A)}.$$

- These calculations show that A^{-1} exists whenever $\det(A) \neq 0$.

The $n \times n$ - case

Definition (Sign of a permutation)

Given a permutation $\sigma : \{1,\ldots,n\} \to \{1,\ldots,n\}$ of n elements, its sign $\mathrm{sgn}(\sigma)$ can be 1 or -1. The sign counts the parity of the number of pairs of elements that are out of order (sometimes called inversions) after applying the permutation. In other words,

$$\mathrm{sgn}(\sigma) = \begin{cases} 1 & \text{if} \quad |(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,n\} \text{ st } i<j,\ \sigma(i) > \sigma(j)| \text{ even} \\ -1 & \text{if} \quad |(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,n\} \text{ st } i<j,\ \sigma(i) > \sigma(j)| \text{ odd} \end{cases}$$

Example

$n = 4$. Consider the permutation π
$\pi(1) = 1$, $\pi(2) = 3$, $\pi(3) = 2$, $\pi(4) = 4$. The pairs (i,j) such that $i < j$ are

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

For all these pairs (i,j) we see $\pi(i) < \pi(j)$ except for $(2,3)$. $\mathrm{sgn}(\pi) = -1$.

The determinant

Definition (Π_n is the set of all permutations of n elements.)

Given $A \in \mathbb{R}^{n \times n}$, the determinant $\det(A)$ is defined as

$$\det(A) = \sum_{\sigma \in \Pi_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} A_{i,\sigma(i)}.$$

Remarks

1. The sign of a permutation is multiplicative, i.e.: for two permutations σ, γ we have that $\operatorname{sgn}(\sigma \circ \gamma) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\gamma)$.
2. For all $n \geq 2$, exactly half of the permutations have sign 1 and exactly half have sign -1.
3. Given a permutation matrix $P \in \mathbb{R}^{n \times n}$ corresponding to a permutation σ, then $\det(P) = \operatorname{sgn}(\sigma)$. We sometimes also write $\operatorname{sgn}(P)$.
4. If A is a 1×1 matrix: there is one permutation of 1 element which has sign 1. It follows $\det(A) = A$.

Further Observations

1. For 2×2 matrices: σ_1 is the identity permutation and σ_2 the permutation that swaps the two elements (which has sign -1).

$$\det(A) = (+1)\prod_{i=1}^{2} A_{i,\sigma_1(i)} + (-1)\prod_{i=1}^{2} A_{i,\sigma_2(i)} = A_{11}A_{22} - A_{12}A_{21}.$$

2. Given a triangular (either upper- or lower-) matrix $T \in \mathbb{R}^{n\times n}$ we have $\det(T) = \prod_{k=1}^{n} T_{kk}$. In particular, $\det(I) = 1$.

Theorem

Given a matrix $A \in \mathbb{R}^{n\times n}$ *we have* $\det(A^\top) = \det(A)$.

Proof.

For a permutation σ let σ^{-1} denote the inverse permutation, i.e.,

$\sigma(i) = j \iff \sigma^{-1}(j) = i$ for all i, j. Note $\mathrm{sgn}(\sigma) = \mathrm{sgn}(\sigma^{-1})$.
$\sum_{\sigma\in\Pi_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} A_{i,\sigma(i)} = \sum_{\sigma^{-1}\in\Pi_n} \mathrm{sgn}(\sigma^{-1}) \prod_{i=1}^{n} A_{\sigma^{-1}(i),i} =$
$\sum_{\sigma\in\Pi_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} A_{\sigma(i),i}$.

General properties of the det-operator

Theorem

- *A matrix* $A \in \mathbb{R}^{n \times n}$ *is invertible if and only if* $\det(A) \neq 0$.
- *Given matrices* $A, B \in \mathbb{R}^{n \times n}$ *we have* $\det(AB) = \det(A)\det(B)$.
- *Given a matrix* $A \in \mathbb{R}^{n \times n}$ *such that* $\det(A) \neq 0$, *then* A *is invertible and*

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Lemma

If $Q \in \mathbb{R}^{n \times n}$ *is an orthogonal matrix then* $\det(Q) = \pm 1$.

Proof.

$1 = \det(I) = \det(Q^\top Q) = \det(Q^\top)\det(Q) = \det(Q)^2$ and so $\det(Q)$ is 1 or -1.

3×3 matrices: there are $3! = 6$ permutations.

$$
\begin{aligned}
\det(A) &= \begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix} \\
&= \begin{vmatrix} A_{11} & & \\ & A_{22} & \\ & & A_{33} \end{vmatrix} + \begin{vmatrix} & A_{12} & \\ A_{21} & & \\ & & A_{33} \end{vmatrix} + \begin{vmatrix} & A_{12} & \\ & & A_{23} \\ A_{31} & & \end{vmatrix} \\
&\quad + \begin{vmatrix} & & A_{13} \\ & A_{22} & \\ A_{31} & & \end{vmatrix} + \begin{vmatrix} & & A_{13} \\ A_{21} & & \\ & A_{32} & \end{vmatrix} + \begin{vmatrix} A_{11} & & \\ & & A_{23} \\ & A_{32} & \end{vmatrix} \\
&= A_{11}A_{22}A_{33} - A_{12}A_{21}A_{33} + A_{12}A_{23}A_{31} \\
&\quad - A_{13}A_{22}A_{31} + A_{13}A_{21}A_{32} - A_{11}A_{23}A_{32}.
\end{aligned}
$$

3×3 matrices: there are $3! = 6$ permutations.

$$
\begin{aligned}
\det(A) &= \begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix} \\
&= \begin{vmatrix} A_{11} & & \\ & A_{22} & \\ & & A_{33} \end{vmatrix} + \begin{vmatrix} & A_{12} & \\ A_{21} & & \\ & & A_{33} \end{vmatrix} + \begin{vmatrix} & A_{12} & \\ & & A_{23} \\ A_{31} & & \end{vmatrix} \\
&\quad + \begin{vmatrix} & & A_{13} \\ & A_{22} & \\ A_{31} & & \end{vmatrix} + \begin{vmatrix} & & A_{13} \\ A_{21} & & \\ & A_{32} & \end{vmatrix} + \begin{vmatrix} A_{11} & & \\ & & A_{23} \\ & A_{32} & \end{vmatrix} \\
&= A_{11}A_{22}A_{33} - A_{12}A_{21}A_{33} + A_{12}A_{23}A_{31} \\
&\quad - A_{13}A_{22}A_{31} + A_{13}A_{21}A_{32} - A_{11}A_{23}A_{32}.
\end{aligned}
$$

There is another convenient way of writing this determinant

$$
\begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix} = A_{11} \begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} - A_{12} \begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{vmatrix} + A_{13} \begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix}.
$$

Co-factors

Definition

Given $A \in \mathbb{R}^{n \times n}$, for each $1 \leq i, j \leq n$ let $\mathscr{A}_{ij}$ denote the $(n-1) \times (n-1)$ matrix obtained by removing row i and column j from A. The co-factors of A are

$$C_{ij} = (-1)^{i+j} \det(\mathscr{A}_{ij}).$$

Lemma

Let $A \in \mathbb{R}^{n \times n}$, for any $1 \leq i \leq n$, $\det(A) = \sum_{j=1}^{n} A_{ij} C_{ij}$.

Lemma

- *The formula we derived for the inverse of 2×2 matrices generalizes:*
- *Given $A \in \mathbb{R}^{n \times n}$ with $\det(A) \neq 0$. Let C be the $n \times n$ matrix with the co-factors of A as entries. We have $A^{-1} = \frac{1}{\det(A)} C^\top$.*
- *One good way to think of this is $AC^\top = \det(A) I$.*

Cramer's Rule: a formula for linear systems

Example $n = 3$. Assume A is n by n and $\det(A) \neq 0$

If $\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$, then we have

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} x_1 & 0 & 0 \\ x_2 & 1 & 0 \\ x_3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} b_1 & A_{12} & A_{13} \\ b_2 & A_{22} & A_{23} \\ b_3 & A_{32} & A_{33} \end{bmatrix}.$$

The determinant is multiplicative

and the determinant of the second matrix in the expression is x_1, i.e., we get

$$\det(A)x_1 = \det(\mathscr{B}_1),$$

where $\mathscr{B}_1$ is the matrix obtained from A by replacing its first column by b. This applies to any any of the columns of A and hence, $x_j = \det(\mathscr{B}_j)/\det(A)$.

Theorem (Cramer's Rule)

Let $A \in \mathbb{R}^{n \times n}$ such that $\det(A) \neq 0$ and $b \in \mathbb{R}^n$ then the solution $x \in \mathbb{R}^n$ of $Ax = b$ is given by

$$x_j = \frac{\det(\mathscr{B}_j)}{\det(A)},$$

where $\mathscr{B}_j$ is the matrix obtained from A by replacing its j-th column by b.

Lemma

The determinant is linear in each row (or each column). In other words, for any $a_0, a_1, a_2 \ldots, a_n \in \mathbb{R}^n$ and $\alpha_0, \alpha_1 \in \mathbb{R}$ we have

$$\begin{vmatrix} - & \alpha_0 a_0^\top + \alpha_1 a_1^\top & - \\ - & a_2^\top & - \\ & \vdots & \\ - & a_n^\top & - \end{vmatrix} = \alpha_0 \begin{vmatrix} - & a_0^\top & - \\ - & a_2^\top & - \\ & \vdots & \\ - & a_n^\top & - \end{vmatrix} + \alpha_1 \begin{vmatrix} - & a_1^\top & - \\ - & a_2^\top & - \\ & \vdots & \\ - & a_n^\top & - \end{vmatrix},$$

and symmetrically for the columns.