WIKIPEDIA

Jacobi's formula

In matrix calculus, **Jacobi's formula** expresses the derivative of the determinant of a matrix A in terms of the adjugate of A and the derivative of A.[1]

If A is a differentiable map from the real numbers to $n \times n$ matrices,

$$\frac{d}{dt} \det A(t) = \operatorname{tr}\left(\operatorname{adj}(A(t)) \, \frac{dA(t)}{dt}\right)$$

where $\operatorname{tr}(X)$ is the trace of the matrix X.

As a special case,

$$\frac{\partial \det(A)}{\partial A_{ij}} = \operatorname{adj}^{\mathrm{T}}(A)_{ij}.$$

Equivalently, if dA stands for the differential of A, the general formula is

$$d \det(A) = \operatorname{tr}(\operatorname{adj}(A) \, dA).$$

It is named after the mathematician Carl Gustav Jacob Jacobi.

Contents

Derivation

Via Matrix Computation

We first prove a preliminary lemma:

Lemma. Let A and B be a pair of square matrices of the same dimension n. Then

$$\sum_i \sum_j A_{ij} B_{ij} = \operatorname{tr}(A^{\mathrm{T}} B).$$

Proof. The product AB of the pair of matrices has components

$$(AB)_{jk} = \sum_i A_{ji} B_{ik}.$$

Replacing the matrix A by its transpose A^{T} is equivalent to permuting the indices of its components:

$$(A^{\mathrm{T}} B)_{jk} = \sum_i A_{ij} B_{ik}.$$

The result follows by taking the trace of both sides:

$$\operatorname{tr}(A^{\mathrm{T}} B) = \sum_j (A^{\mathrm{T}} B)_{jj} = \sum_j \sum_i A_{ij} B_{ij} = \sum_i \sum_j A_{ij} B_{ij}. \ \square$$

Theorem. (Jacobi's formula) For any differentiable map A from the real numbers to $n \times n$ matrices,

$$d \det(A) = \operatorname{tr}(\operatorname{adj}(A)\, dA).$$

Proof. Laplace's formula for the determinant of a matrix A can be stated as

$$\det(A) = \sum_j A_{ij} \operatorname{adj}^{\mathrm{T}}(A)_{ij}.$$

Notice that the summation is performed over some arbitrary row i of the matrix.

The determinant of A can be considered to be a function of the elements of A:

$$\det(A) = F\,(A_{11}, A_{12}, \ldots, A_{21}, A_{22}, \ldots, A_{nn})$$

so that, by the chain rule, its differential is

$$d \det(A) = \sum_i \sum_j \frac{\partial F}{\partial A_{ij}}\, dA_{ij}.$$

This summation is performed over all $n \times n$ elements of the matrix.

To find $\partial F / \partial A_{ij}$ consider that on the right hand side of Laplace's formula, the index i can be chosen at will. (In order to optimize calculations: Any other choice would eventually yield the same result, but it could be much harder). In particular, it can be chosen to match the first index of $\partial \,/\, \partial A_{ij}$:

$$\frac{\partial \det(A)}{\partial A_{ij}} = \frac{\partial \sum_k A_{ik} \operatorname{adj}^{\mathrm{T}}(A)_{ik}}{\partial A_{ij}} = \sum_k \frac{\partial (A_{ik} \operatorname{adj}^{\mathrm{T}}(A)_{ik})}{\partial A_{ij}}$$

Thus, by the product rule,

$$\frac{\partial \det(A)}{\partial A_{ij}} = \sum_k \frac{\partial A_{ik}}{\partial A_{ij}} \operatorname{adj}^{\mathrm{T}}(A)_{ik} + \sum_k A_{ik} \frac{\partial \operatorname{adj}^{\mathrm{T}}(A)_{ik}}{\partial A_{ij}}.$$

Now, if an element of a matrix A_{ij} and a cofactor $\operatorname{adj}^{\mathrm{T}}(A)_{ik}$ of element A_{ik} lie on the same row (or column), then the cofactor will not be a function of A_{ij}, because the cofactor of A_{ik} is expressed in terms of elements not in its own row (nor column). Thus,

$$\frac{\partial \operatorname{adj}^{\mathrm{T}}(A)_{ik}}{\partial A_{ij}} = 0,$$

so

$$\frac{\partial \det(A)}{\partial A_{ij}} = \sum_k \operatorname{adj}^{\mathrm{T}}(A)_{ik} \frac{\partial A_{ik}}{\partial A_{ij}}.$$

All the elements of A are independent of each other, i.e.

$$\frac{\partial A_{ik}}{\partial A_{ij}} = \delta_{jk},$$

where δ is the Kronecker delta, so

$$\frac{\partial \det(A)}{\partial A_{ij}} = \sum_k \operatorname{adj}^{\mathrm{T}}(A)_{ik} \delta_{jk} = \operatorname{adj}^{\mathrm{T}}(A)_{ij}.$$

Therefore,

$$d(\det(A)) = \sum_i \sum_j \operatorname{adj}^{\mathrm{T}}(A)_{ij} \, dA_{ij},$$

and applying the Lemma yields

$$d(\det(A)) = \operatorname{tr}(\operatorname{adj}(A) \, dA). \ \square$$

Via Chain Rule

Lemma 1. $\det'(I) = \operatorname{tr}$, where $\det'$ is the differential of $\det$.

This equation means that the differential of $\det$, evaluated at the identity matrix, is equal to the trace. The differential $\det'(I)$ is a linear operator that maps an $n \times n$ matrix to a real number.

Proof. By definition of the differential, we have

$$\det'(I)(T) = \lim_{\varepsilon \to 0} \frac{\det(I + \varepsilon T) - \det I}{\varepsilon}$$

$\det(I + \varepsilon T)$ is a polynomial in ε of order n. It is closely related to the characteristic polynomial of T. The constant term ($\varepsilon = 0$) is 1, while the linear term in ε is $\operatorname{tr} T$.

Lemma 2. For an invertible matrix A, we have: $\det'(A)(T) = \det A \operatorname{tr}(A^{-1}T)$.

Proof. Consider the following function of X:

$$\det X = \det(AA^{-1}X) = (\det A)\ \det(A^{-1}X)$$

We calculate the differential of $\det X$ and evaluate it at $X = A$ using Lemma 1, the equation above, and the chain rule:

$$\det'(A)(T) = \det A\ \det'(I)(A^{-1}T) = \det A \operatorname{tr}(A^{-1}T)$$

Theorem. (Jacobi's formula) $\frac{d}{dt} \det A = \operatorname{tr}\left(\operatorname{adj} A \frac{dA}{dt}\right)$

Proof. If A is invertible, by Lemma 2, with $T = dA/dt$

$$\frac{d}{dt} \det A = \det A \operatorname{tr}\left(A^{-1}\frac{dA}{dt}\right) = \operatorname{tr}\left(\operatorname{adj} A\ \frac{dA}{dt}\right)$$

using the equation relating the adjugate of A to A^{-1}. Now, the formula holds for all matrices, since the set of invertible linear matrices is dense in the space of matrices.

Corollary

The following is a useful relation connecting the trace to the determinant of the associated matrix exponential:

$$\boxed{\det e^{tB} = e^{\operatorname{tr}(tB)}}$$

This statement is clear for diagonal matrices, and a proof of the general claim follows.

For any invertible matrix $A(t)$, in the previous section "Via Chain Rule", we showed that

$$\frac{d}{dt} \det A(t) = \det A(t)\ \operatorname{tr}\left(A(t)^{-1}\ \frac{d}{dt}A(t)\right)$$

Considering $A(t) = \exp(tB)$ in this equation yields:

$$\frac{d}{dt} \det e^{tB} = \operatorname{tr}(B) \det e^{tB}$$

The desired result follows as the solution to this ordinary differential equation.

Applications

Several forms of the formula underlie the Faddeev–LeVerrier algorithm for computing the characteristic polynomial, and explicit applications of the Cayley–Hamilton theorem. For example, starting from the following equation, which was proved above:

$$\frac{d}{dt} \det A(t) = \det A(t) \ \operatorname{tr}\left(A(t)^{-1} \, \frac{d}{dt} A(t)\right)$$

and using $A(t) = tI - B$, we get:

$$\frac{d}{dt} \det(tI - B) = \det(tI - B) \operatorname{tr}[(tI - B)^{-1}] = \operatorname{tr}[\operatorname{adj}(tI - B)]$$

where adj denotes the adjugate matrix.

Remarks

1. Magnus & Neudecker (1999), Part Three, Section 8.3

References

- Magnus, Jan R.; Neudecker, Heinz (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics* (https://books.google.com/books?id=0CXXdKKiIpQC). Wiley. ISBN 0-471-98633-X.
- Bellmann, Richard (1997). *Introduction to Matrix Analysis* (https://books.google.com/books?id=QVCflvTPYE8C). SIAM. ISBN 0-89871-399-4.