A geometric approach to the Cauchy-Binet formula

Gopala Krishna Srinivasan

Abstract

The Cauchy-Binet formula is one of the most important and substantially non-trivial result on the theory of determinants. We provide an interesting geometric proof of this important result obtaining it as a corollary of a new proof of the formulas for the volume of a k-parallelpiped in $n-$dimensional space.

Mathematics Subject Classification 15A15 15A18 15A24
Keywords and phrases : Cauchy-Binet Formula, Volumes of $k-$ parallelpipeds, Gram-determinant

§1 Introduction

The Cauchy-Binet formula asserts that if A is a $m \times n$ matrix and B is an $n \times m$ matrix where $m \leq n$, then

$$\text{Det}(AB) = \text{sum of the principal } m \times m \text{ minors of } B^T A^T \tag{1.1}$$

the superscript T denoting the transpose. The formula is of an ancient vintage going back more than two centuries (see [8]). The case $m = 2$ which reads:

$$\sum_{i<j} (a_i b_j - a_j b_i)^2 = \|a\|^2 \|b\|^2 - (a_1 b_1 + \ldots + a_n b_n)^2, \tag{1.2}$$

(the norm being the Euclidean norm) is perhaps even older and is sometimes known as the *"Lagrange identity"*.

Despite its importance, the result, with a few exceptions is conspicuously absent in most linear algebra books except ones a slant towards combinatioral applications such as Marcus and Minc [7]. Among the exceptional books containing an account of it are the classic works of [2], [6] and the recent and important book of Denis Serre [9] - rich in analytic content. Thus it is not surprising that the result has resurfaced often with remarkable proofs many with a combinatorial flavor. In the recent paper [1] the author has given a proof of the result with numerous references and some further ramifications as well. The expression on the right hand side of (1.1) should suggest an approach through determining eigen-values of a certain matrix which is the approach followed here.

A closely related issue is the formula for the volume of the $k-$parallelpiped spanned by k vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ in $\mathbb{R}^n$. The volume is defined inductively as follows. Denoting by $\mathrm{Vol}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{k-1})$ the volume of the $(k-1)$ parellelpiped spanned by the first $k-1$ vectors, we define

$$\mathrm{Vol}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k) = \|\mathbf{q}_k\| \mathrm{Vol}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{k-1}) \tag{1.3}$$

where $\mathbf{q}_k = \mathbf{v}_k - \mathbf{p}_k$ and $\mathbf{p}_k$ is the orthogonal projection of $\mathbf{v}_k$ on the linear subspace spanned by $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{k-1}$. It is not apriori clear that the volume so defined is independent of the ordering of the vectors but it follows as a consequence of the formula for the volume in terms of the Gram determinant or *Grammian* namely,

$$G(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \mathrm{Det}(\langle \mathbf{v}_i, \mathbf{v}_j \rangle) \tag{1.4}$$

It follows at once from (1.1) that

$$\mathrm{Vol}(\mathbf{v}_1, \ldots, \mathbf{v}_k) = G(\mathbf{v}_1, \ldots, \mathbf{v}_k) := \text{sum of squares of the principal } m \times m \text{ minors of } M^T M \tag{1.5}$$

where M denotes the $n \times k$ matrix given by

$$M = [\mathbf{v}_1, \ldots, \mathbf{v}_k]$$

The right hand side of (1.5) is usually denoted by

$$\|\mathbf{v}_1 \wedge \mathbf{v}_2 \wedge \ldots \wedge \mathbf{v}_k\|^2 \tag{1.6}$$

The expression (1.5) is "quadratic" in nature and arises by taking $A = B = M$ in (1.1). The latter being "bilinear" in nature, we could recover (1.1) from (1.5) through polarization. Thus it suffices to prove (1.5) and in this paper we do this geometrically thereby providing yet another proof of the Cauchy-Binet formula. In [5] the author has discussed (1.5) in the light of singular value decomposition of M and writes the volume as the product of the singular values. For completeness we also provide a proof (with minimal details) that the volume of the $k-$parallelpiped is the square root of the Gram determinant.

The importance of (1.5) in analysis becomes transparent if we recall the definition of the $k-$volume of an immersed $k-$ dimensional manifold in $\mathbb{R}^n$, given parametrically by a smooth map

$$\Phi \; : \; I \longrightarrow \mathbb{R}^n, \quad (t_1, t_2, \ldots, t_k) \mapsto \Phi(t_1, t_2, \ldots, t_k), \tag{1.7}$$

of rank k where I is an open set in $\mathbb{R}^k$. The $k-$ volume is given by

$$\int_I \sqrt{G(\Phi_{t_1}, \Phi_{t_2}, \ldots, \Phi_{t_k})} dt_1 dt_2 \ldots dt_k. \tag{1.8}$$

It is of immediate concern to decide whether the formula given by (1.8) is invariant under a diffeomorphic change of parametrization:

$$\psi : J \longrightarrow I, \quad (s_1, \ldots, s_k) \mapsto (t_1, \ldots, t_k) = \psi(s_1, \ldots, s_k)$$

This is at once clear if we use the expression (1.6) since

$$\|\Phi_{t_1} \wedge \Phi_{t_2} \wedge \ldots \wedge \Phi_{t_k}\|^2 = \|\Psi_{t_1} \wedge \Psi_{t_2} \wedge \ldots \wedge \Psi_{t_k}\|^2 \Big(\mathrm{Det}\Big(\frac{\partial t_j}{\partial s_i}\Big)\Big)^2$$

whereas the use of the alternate expression given by the Grammian has the disadvantage of having to work in the ambiant space $\mathbb{R}^n$ in which the $k-$dimensional manifold is immersed. Formulas (1.1) and (1.5) could be of potential use in the study of higher curvatures of space curves [3], [4].

§2 Proofs of the results:

For completeness we briefly indicate the proof that the volume of the $k-$ parallelpided is given by (1.4)

Theorem 2.1. The volume of the $k-$parallelpiped spanned by $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ equals the Gram determinant $G(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k) := \mathrm{Det}(\langle \mathbf{v}_i, \mathbf{v}_j \rangle)$. In particular it is independent of the ordering of the vectors. Also, the Gram determinant is non-negative.

Proof: Using the notation explained in the introduction, we show that

$$\|\mathbf{q}_j\|^2 = G(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_j)/G(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{j-1}), \quad j = 2, 3, \ldots, k, \tag{2.1}$$

This is clear for $j = 2$. Proceeding by induction on j, we have the following system of equations characterizing $\mathbf{q}_j$:

$$\mathbf{q}_j = \mathbf{v}_j - (c_1\mathbf{v_1} + c_2\mathbf{v}_2 + \ldots + c_{j-1}\mathbf{v}_{j-1}), \tag{2.2}$$

where the coefficients are subject to the constraints:

$$\langle \mathbf{q}_j, \mathbf{v}_i \rangle = 0, \quad i = 1, 2, \ldots, j-1, \tag{2.3}$$

Taking the dot product of (2.2) with $\mathbf{v}_i$ at once leads to a system of linear equations for the coefficients which can be determined by Crammer's rule. Finally using the obvious fact

$$\|\mathbf{q}_j\|^2 = \langle \mathbf{v}_j, \mathbf{q}_j \rangle$$

we quickly get by using the expression for the c_i obtained via Crammer's rule, we obtain (2.1) and the theorem follows from the inductive definition (1.3).

We now prove the important formula (1.5).

Theorem 2.2 For vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in \mathbb{R}^n$, we have

$$G(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k) = H(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k) \tag{2.4}$$

where we have used the notation $H(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ to denote the right hand side of (1.5).

Proof: Observe that both the functions G and H satisfy the following properties

(i) They are both invariant under permutation of the vectors.

(ii) They both vanish if the set of vectors are linearly dependent. For the Grammian this follows from theorem (2.1).

(iii) If $\mathbf{v}_j$ is replaced by $c\mathbf{v}_j$ both get scaled by a factor of c^2.

(iv) If $\mathbf{v}_j$ is replaced by $\mathbf{v}_j + c\mathbf{v}_i$ with $i \neq j$ then both G and H remain invariant.

We proceed by assuming that the vectors are linearly independent and

$$H(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = H(\|\mathbf{v}_1\|\mathbf{u}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = \|\mathbf{v}_1\|^2 H(\mathbf{u}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = G(\mathbf{v_1})H(\mathbf{u}_1, \mathbf{v}_2, \dots, \mathbf{v}_k).$$

where $\mathbf{u}$ is the unit vector $\mathbf{v}_1/\|\mathbf{v}_1\|$. Replace $\mathbf{v}_2$ by $\mathbf{q}_2 = \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{u}_1 \rangle \mathbf{u}_1$ and we get using (iv), (iii) and equation (2.1) that

$$H(\mathbf{u}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = G(\mathbf{v_1})\|\mathbf{q}_2\|^2 H(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{v}_k) = G(\mathbf{v}_1, \mathbf{v}_2) H(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{v}_k).$$

where, $\mathbf{u}_2 = \mathbf{q}_2/\|\mathbf{q}_2\|$. Proceeding thus we finally get

$$H(\mathbf{u}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = G(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) H(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k). \tag{2.5}$$

where $\{\mathbf{u}_1, \dots, \mathbf{u}_k)$ is an orthonormal $k-$frame in $\mathbb{R}^n$. To complete the proof we merely have to show that

$$H(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k) = 1. \tag{2.6}$$

The symmetric $n \times n$ matrix

$$A = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k][\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k]^T$$

has rank k and so zero is an eigen-value with geometric (and hence algebraic) multiplicity $n-k$. It is also clear that the vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ are all eigen-vectors of A with eigen-value one. Thus one is an eigen-value with geometric (and hence algebraic) multiplicity k. Thus the $k-$th symmetric function of the eigen-values of A is one which in turn is the sum of the $k \times k$ principal minors of A and in this in turn agrees with the sum of the squares of all the $k \times k$ minors of $[\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k]$. Thus (2.6) has been established and the proof of the theorem is complete.

Corollary 2.3 (Cauchy-Binet): if A is a $m \times n$ matrix and B is an $n \times m$ matrix where $m \leq n$, then

$$\text{Det}(AB) = \text{sum of squares of the principal } m \times m \text{ minors of } B^T A^T$$

Proof: Successive polarization of the quadratic form to the associated bilinear form in the variables $\mathbf{v}_1, \dots, \mathbf{v}_k$ (taken one at a time) would produce the result.

Gopala Krishna Srinivasan
Department of Mathematics,
Indian Institute of Technology Bombay
Powai, Mumbai 400 076,
INDIA
gopal@math.iitb.ac.in

References

[1] T. Konstantopoulos, *A multilinear algebra proof of the Cauchy Binet formula and a mulit-linear version of Parseval's identity*, Linear algebra and its applications, **439** (2013) 2651-2658.

[2] F. Gantmacher, *The theory of matrices, Volume - I* , AMS Chelsea Publishing, American Mathematical Society, Rhode Island, 2000.

[3] H. Gluck, *Higher curvature of curves in Euclidean space - I*, American Mathematical Monthly, **73** (1966) 966-704

[4] H. Gluck, *Higher curvature of curves in Euclidean space - II*, American Mathematical Monthly, **74** (1967) 1049-1056.

[5] A. B-Israel, *A volume associated with $m \times n$ matrices* Linear algebra and its applications, **167** (1992) 87-111

[6] P. Lancaster and M. Tismenetsky, *Theory of matrices, Second Edition*, Academic Press, New York, 1985.

[7] M. Marcus and H. Minc, *Introduction to linear algebra*, Dover Publication, New York, 1988.

[8] T. Muir, *The theory of determinant in the historical order of its development, Part - I, Determinants in general Leibnitz (1693) to Cayley (1841)*, Macmillan and Company, London, 1890.

[9] D. Serre, *Matrices, Theory and applications* Graduate Texts in Mathematics 216, Springer Verlag, New York (2002).