Subdiscriminants of symmetric matrices are sums of squares

Marie-Françoise Roy*

February 6, 2005

In this Note we prove that the subdiscriminants of a symmetric matrix are sums of squares. This generalizes a result of [2] stating that the discriminant of a symmetric is a sum of squares and is inspired by its proof. A different, less explicit proof that the discriminant of a symmetric is a sum of squares also apear in [3]. As a consequence, we obtain an algebraic proof of the fact that all the roots of the characteristic polynomial of a symmetric matrix are real.

1 Discriminants and subdiscriminants

We consider an ordered field K, a real closed field R containing K and $\mathrm{C} = \mathrm{R}[i]$ its algebraic closure.

The k-subdiscriminant of a monic polynomial $P \in \mathrm{K}[X]$ of degree p is defined as follows. Let $x_1, \ldots, x_p$ be the multiset of roots of P in C, counted with multiplicities. The k-subdiscriminant of P , $0 \le j \le p-1$, is by definition

$$\mathrm{SubDisc}_k(P) = \sum_{I \subset \{1,\ldots,p\}, \#(I)=p-k} \Pi_{(j,\ell)\in I, \ell > j} (x_j - x_\ell)^2.$$

Note that $\mathrm{SubDisc}_{p-1}(P) = p$. Subdiscriminants generalize the classical notion of discriminant which is the 0-th subdiscriminant:

$$\mathrm{SubDisc}_0(P) = \mathrm{Disc}(P) = \Pi_{p \ge j > \ell \ge 1} (x_j - x_\ell)^2.$$

Subdiscriminants turn out to be determinants of matrices with entries the Newton nums of P. Denoting by $\mathrm{N}_i(P)$ the Newton sum $\sum_{i=1,\ldots,p} x_i^j$, let the Hermite matrix $\mathrm{Her}_k(P)$ be the $(p-k) \times (p-k)$-matrix with i, j-th entry $\mathrm{N}_{i+j-2}(P)$ $i, j = 1, \ldots p-k$.

Proposition 1

$$\mathrm{SubDisc}_k(P) = \det(\mathrm{Her}_k(P))$$

The proof uses the classical Cauchy-Binet formula

*IRMAR (URA CNRS 305), Université de Rennes, Campus de Beaulieu 35042 Rennes cedex FRANCE, email: marie-francoise.roy@univ-rennes1.fr

Proposition 2 *Let A be a $n \times m$ matrix and B be a $m \times n$ matrix. For every $I \subset \{1,,m\}$ of cardinality n, denote by A_I the $n \times n$ matrix obtained by extracting from A the columns with indices in I. Similarly let B^I be the $n \times n$ matrix obtained by extracting from B the rows with indices in I.*

$$\det(AB) = \sum_{I \subset \{1,,m\}, \#(I)=n} \det(A_I) \det(B^I).$$

Proof of Proposition 1: Define

$$V_k = \begin{bmatrix} 1 & 1 & . & . & . & . & 1 \\ x_1 & x_2 & . & . & . & . & x_p \\ . & . & . & . & . & . & . \\ x_1^{p-k-1} & x_2^{k-1} & . & . & . & . & x_p^{p-k-1} \end{bmatrix}.$$

It is clear that $V_k V_k^t = \mathrm{Her}_k(P)$. Now apply Binet-Cauchy formula, noting that, if $I \subset \{1,,p\}, \#(I) = p-k$

$$\det(V_{kI}) = \Pi_{(j,\ell)\in I, \ell > j}(x_j - x_\ell).$$

□

Let A is a symmetric $p \times p$ matrix with coefficients in a ring A. We define the k-th subdiscriminant of A as the determinant of the matrix $\mathrm{Her}_k(A)$ whose (i,j)-th entry is $\mathrm{Tr}(A^{i+j-2})$, $i,j = 1, \ldots p-k$. When A is with entries in a field K, the k-th subdiscriminant of A coincides with the k-th subdiscriminant of its characteristic polynomial P. Indeed, the Newton sum $\mathrm{N}_i(P)$ of A is $\mathrm{Tr}(A^i)$, the trace of the matrix A^i.

2 Orthogonal basis of symmetric matrices

We define a linear basis $E_{j,\ell}$ of the space $\mathrm{Sym}(p)$ of symmetric matrices of size p as follows. First define $F_{j,\ell}$ as the matrix having all zero entries except 1 at (j,ℓ). Then take $E_{j,j} = F_{j,j}, E_{j,\ell} = 1/\sqrt{2}(F_{j,\ell} + F_{\ell,j}), \ell > j$. Define E as the ordered set $E_{j,\ell}$ $p \geq \ell \geq j \geq 0$, indices being taken in the order

$$(1,1), \ldots, (p,p), (1,2), \ldots, (1,p), \ldots, (p-1,p).$$

For simplicity, we index elements of E pairs (j,ℓ), ℓ.

It is immediate to check that the map associating to $(A,B) \in \mathrm{Sym}(p) \times \mathrm{Sym}(p)$ the value $\mathrm{Tr}(AB)$ is a scalar product on $\mathrm{Sym}(p)$ with orthogonal basis E.

Let B_k be the $(p-k) \times p(p+1)/2$ matrix with $(i,(j,\ell))$-th entry the (j,ℓ)-th component of A^{i-1} in the basis E.

Proposition 3

$$\mathrm{Her}_k(A) = B_k \times B_k^t.$$

Proof : Immediate since $\mathrm{Tr}(A^{i+j})$ is the scalar product of A^i by A^j in the basis E. □

3 Subdiscriminants of symmetric matrices are sums of squares

We consider a generic symmetric matrix $A = [ai, j]$ whose entries are $p(p+1)/2$ independant variables $a_{j,\ell}, \ell \geq j$. We are going to give an explicit expression of $\mathrm{SubDisc}_k(A)$ as a sum of products of powers of 2 by squares of elements of the ring $\mathbb{Z}[a_{j,\ell}]$.

Let B_k be the $(p-k) \times p(p+1)/2$ matrix with $(i,(j,\ell))$-th entry the (j,ℓ)-th component of A^{i-1} in the basis E.

Proposition 4 $\mathrm{SubDisc}_k(A)$ *is the sum of squares of the* $(p-k)\times(p-k)$ *minors of* B_k.

Proof : Use Proposition 3 and Binet-Cauchy formula. □

Noting that a $(p-k) \times (p-k)$ minor of B_k is a power of 2 multiplied by a square of an element of $\mathbb{Z}[a_{j,\ell}]$, we obtain an explicit expression of $\mathrm{SubDisc}_k(A)$ as a sum of products of powers of 2 by squares of elements of the ring $\mathbb{Z}[a_{j,\ell}]$.

As a consequence the k-th subdiscriminant of a symmetric matrix with coefficients in a ring A is a sum of products of powers of 2 by squares of elements in A.

Let us take a simple example and consider

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}.$$

The characteristic polynomial of A is $X^2 - (a_{11} + a_{22})X + a_{11}a_{22} - a_{12}^2$, and its discriminant is $(a_{11}+a_{22})^2 - 4(a_{11}a_{22} - a_{12}^2)$. On the other hand the sum of the squares of the 2 by 2 minors of

$$B_0 = \begin{bmatrix} 1 & 1 & 0 & \\ a_{11} & a_{22} & \sqrt{2} & a_{22} \end{bmatrix}$$

is

$$(a_{22} - a_{11})^2 + (\sqrt{2}a_{12})^2 + (\sqrt{2}a_{12})^2.$$

It is easy to check the statement of Proposition 4 in this particular case.

4 Characteristic polynomials of symmetric matrices are hyperbolic

By definition, a polynomial $P \in \mathrm{R}[X]$ is hyperbolic if all its roots are in R.

We give an algebraic proof of the classical theorem.

Proposition 5 *The characteristic polynomial of a symmetric matrix is hyperbolic.*

Proof : We denote by P the characteristic polynomial of a matrix A.

First note that, by Proposition 4 $\mathrm{SubDisc}_i(A) = 0$ if only if the rank of B_i is less than $n - i$. It follows that $\mathrm{SubDisc}_k(A) > 0$ implies $\mathrm{SubDisc}_i(A) > 0$ for every $n - 1 \geq i \geq k$, and $\mathrm{SubDisc}_{k-1}(A) = 0$ implies $\mathrm{SubDisc}_i(A) = 0$ for every $0 \leq i < k$. In other words, for every symmetric matrix A, there exists $k, n - 1 \geq k \geq 0$ such that the signs of the subdiscriminants of A are

$$(\wedge_{p-1 \geq i \geq k} \mathrm{SubDisc}_i(A) > 0 \wedge \wedge_{0 \leq i < k} \mathrm{SubDisc}_i(A) = 0).$$

So the number of roots of the characteristic polynomial P of A is $p - k$, using Proposition (relation between subresultants and subdiscriminants) and Proposition (counting numb er of real roots in terms of permanencies minus variations) of [1], while the number of distinct roots of P is $p - k$ using Proposition (subresultants give degree of gcd). □

Since it is clear that every hyperbolic polynomial is the characteristic polynomial of a diagonal symmetric matrix with entries in R, Proposition 5 imples that the set of hyperbolic polynomials is characterized by

$$\vee_{k=p-1,\ldots,0} (\wedge_{p-1 \geq i \geq k} \mathrm{SubDisc}_i(P) > 0 \wedge \wedge_{0 \leq i < k} \mathrm{SubDisc}_i(P) = 0).$$

On the other hand, the sign condition

$$\mathrm{SubDisc}_{p-2}(P) \geq 0 \wedge \ldots \wedge \mathrm{SubDisc}_0(P) \geq 0$$

does not imply that P is hyperbolic: the polynomials $X^4 + 1$ has no real root (its four roots are $\pm\sqrt{2}/2 \pm i\sqrt{2}/2$), and it is immedate to check uthat is satisties $\mathrm{SubDisc}_2(P) = \mathrm{SubDisc}_1(P) = 0, \mathrm{SubDisc}_0(A) > 0$.

In fact, the set of hyperbolic polynomials the closure of the set defined by

$$\mathrm{SubDisc}_{p-2}(P) > 0 \wedge \ldots \wedge \mathrm{SubDisc}_0(P) > 0,$$

but does not coincide with the set defined by

$$\mathrm{SubDisc}_{p-2}(A) \geq 0 \wedge \ldots \wedge \mathrm{SubDisc}_0(A) \geq 0.$$

References

[1] S. Basu, R. Pollack,M.-F. Roy, *Algorithms in real algebraic geometry*, Springer (2003).

[2] N. V. Ilyusheckin, *On some identities for the elements of a symmetric matrix* "Zapiski Nauchnyh Seminarov POMI" Vol. 303 "Investigations on Linear Operators and Function Theory. Part 31" editor S.V.Kislyakov, http://www.pdmi.ras.ru/znsl/2003/v303.html.

[3] P. Lax, *On the discriminant of real symmetric matrices* Communications on pure and applied mathematics Vol. LI 1387-1396 (1998).