On Compositions of Quadratic Forms in Many Variables[1]

Identities like $(x^2+y^2)(z^2+w^2) = (xz-yw)^2 + (xw+yz)^2$ and Lagrange's identity are useful in Number Theory; when do they exist? To answer such a question, it is necessary first to put it in clear terms. We ask for what values of m, n, ℓ does there exist an identity of the form

$$(x_1^2 + x_2^2 + \cdots + x_n^2)(y_1^2 + y_2^2 + \cdots + y_m^2) = (z_1^2 + z_2^2 + \cdots + z_\ell^2),$$

where each z_i is a bilinear function of the $m+n$ variables x_1 up to x_n and y_1 up to y_m. What is a bilinear function (or combination) and why do we require it? It is simply a sum of the form $\sum c_{ij}x_iy_j$. This is exactly the form of the known identities, and is also very natural. In this paper we are only concerned in the case $n = m = \ell$. There are known identities for $n = 1$, 2, 4, 8. The identity for $n = 1$ is simply $x^2y^2 = (xy)^2$, the identity for $n = 2$ was listed above, the identity for $n = 4$ is Lagrange's identity, and the identity for $n = 8$ is too long and cumbersome to write down here[2]. We note that the identity for $n = 1$ comes from real multiplication, the one for $n = 2$ comes from complex multiplication, the one for $n = 4$ comes from quaternion multiplication, and the one for $n = 8$ comes from octonion multiplication[3]. In fact, in some sense, these are the only possible identities (we do not pursue this here).

How can we represent such a composition formula? In other words, how do we represent the bilinear combinations forming the z_is? We shall use an $n \times n$ matrix whose entries are linear functions of the x_is (functions of the form $\sum c_i x_i$). Each row represents a different bilinear combination. Denoting by a_{ij} the element at row i and column j, we have $z_i = \sum a_{ij}y_j$. When does such a matrix A represent a composition formula? Let us compute:

$$\sum_{i=1}^{n} z_i^2 = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij}y_j \right)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{ij}a_{ik}y_jy_k.$$

The transpose A' of a matrix is obtained by interchanging the rows and columns. It is easy to see that the sum $\sum_i a_{ij}a_{ik}$ is the (j,k)th element of the matrix AA'. Our required composition formula can be written

$$\sum_{i=1}^{n} z_i^2 = \sum_{j=1}^{n}\sum_{k=1}^{n} x_j^2 y_k^2$$

and we see that if $j = k$ then the (j,k)th element of AA' should equal $\sum x_i^2$, and otherwise it should equal zero. In other words, $AA' = \sum x_i^2$ (this is a shorthand for $(\sum x_i^2)I$ where I is the identity matrix of appropriate size).

Now let us find out when does the equality $AA' = \sum x_i^2$ hold. Recall that each element of A is actually a linear function of the x_is, so we can write $A = \sum x_i A_i$, where each A_i is an ordinary matrix. This turns the required equality to the form $(\sum x_i A_i)(\sum x_i A_i') = \sum x_i^2$, using properties of the transpose. The next step is a further reduction. Let us single out the last variable x_n. Substituting $x_1 = x_2 = \cdots = x_{n-1} = 0$ and $x_n = 1$, we see that $A_nA_n' = I$, and so A_n and A_n' are inverses (such matrices are called orthogonal). This means in particular that $A_n'A_n = I$.

---

[1] Based on the paper "Über die Komposition der quadratischen Formen von beliebig vielen Variablen" by Adolf Hurwitz, number LXXXII in his collected works.

[2] I'm using here the words of Herstein. Hurwitz in fact listed this identity.

[3] The octonions are non-associative (and also non-commutative, like the quaternions), and form a vector space of dimension 8 over the reals.

Now let $B_i = A_i A_n'$. Note that $B_n = I$ and that $B_i' = A_n A_i'$. The basic equality satisfied by these new matrices is the following:

$$\sum_{i=1}^{n} x_i^2 = \left(\sum_{i=1}^{n} x_i A_i\right)\left(\sum_{i=1}^{n} x_i A_i'\right) = \left(\sum_{i=1}^{n} x_i A_i\right) A_n' A_n \left(\sum_{i=1}^{n} x_i A_i'\right)$$
$$= \left(\sum_{i=1}^{n} x_i A_i A_n'\right)\left(\sum_{i=1}^{n} x_i A_n A_i'\right) = \left(\sum_{i=1}^{n} x_i B_i\right)\left(\sum_{i=1}^{n} x_i B_i'\right).$$

If we expand the righthand side, we see that the coefficient of x_i^2 is $B_i B_i'$, whereas the coefficient of $x_i x_j$ (for $i \neq j$) is $B_i B_j' + B_j B_i'$. Since the identity should be true if arbitrary numbers are substituted for the x_is, it follows that $B_i^2 = I$ and $B_i B_j' = -B_j B_i'$ for $i \neq j$. In particular, for $j = n$ we get $B_i = -B_i'$. We call such matrices skew-symmetric (matrices like AA' which satisfy $M = M'$ are called symmetric). It also follows that $B_i^2 = -B_i B_i' = -I$. When is this possible? Since $\det B_i^2 = (\det B_i)^2$ and $\det I_n = (-1)^n$, it follows that n must be even (unless $n = 1$, in which case the only equality available to us is $B_1 B_1' = I$) and $\det B_i^2 = 1$. In particular, each B_i is regular (also called non-singular and invertible). Another use of the skew-symmetry is noting that $B_i B_j = -B_i B_j' = B_j B_i' = -B_j B_i$.

Now let us consider the 2^{n-1} possible products of different matrices from the set $\{B_1, \ldots, B_{n-1}\}$. Such products are very special: they are all either symmetric or skew-symmetric, depending on the number of factors. We now show why this is so. Consider the product $B_{i_1} B_{i_2} \cdots B_{i_m}$ containing m factors. Taking the transpose, the order of the factors is reversed:

$$(B_{i_1} B_{i_2} \cdots B_{i_m})' = B_{i_m}' B_{i_{m-1}}' \cdots B_{i_1}'.$$

Since all the B_is are skew-symmetric, we can get rid of the transposes:

$$B_{i_m}' B_{i_{m-1}}' \cdots B_{i_1}' = (-1)^m B_{i_m} B_{i_{m-1}} \cdots B_{i_1}.$$

Next, let us try to restore the order of the matrices. Bringing B_{i_1} to the front requires $m - 1$ transpositions (operations of the type $B_i B_j = -B_j B_i$), which incurs an overhead of $(-1)^{m-1}$. Next, bringing B_{i_2} to lie just beside it requires $m - 2$ transpositions. Continuing this way, we get that

$$(B_{i_1} B_{i_2} \cdots B_{i_m})' = (-1)^{m+(m-1)+\cdots+1} B_{i_1} B_{i_2} \cdots B_{i_m}.$$

Now $m + (m-1) + \cdots + 1 = m(m+1)/2$, which is even if $m \equiv 0$ or $3 \pmod 4$, and odd otherwise. This is so because $m(m+1)/2$ is odd or even depending on the residue modulo 4 of $m(m+1)$, and this reduces to four trivial cases. Hence, our product is symmetric when $m \equiv 0$ or $3 \pmod 4$, and skew-symmetric otherwise.

Now we wish to find out what is the rank of these 2^{n-1} matrices, that is how many of them are linearly independent. For this, consider a linear dependency containing I, the empty product (a linear dependency is a linear combination which sums to 0; we say that a linear combination contains a matrix if its coefficient is non-zero). We can write any such combination in the form $I = L$, where L is some linear combination of the other matrices. Let us separate the combination L into symmetric and skew-symmetric matrices, so that $L = E + O$, where E is a linear combination of symmetric matrices and O is a linear combination of skew-symmetric matrices. Since I is symmetric we have $E - O = E' + O' = L' = L = E + O$, and so O sums to zero. Hence we can assume that L is a linear combination of symmetric matrices (this means that from the existence of any linear combination L it follows the existence of one containing no skew-symmetric matrices). Such a linear combination looks like this:

$$I = \sum_{i,j,k} c_{i,j,k} B_i B_j B_k + \sum_{i,j,k,\ell} c_{i,j,k,\ell} B_i B_j B_k B_\ell + \cdots.$$

Suppose we multiply this by some B_m on the right:

$$B_m = \sum_{i,j,k} c_{i,j,k} B_i B_j B_k B_m + \sum_{i,j,k,\ell} c_{i,j,k,\ell} B_i B_j B_k B_\ell B_m + \cdots .$$

As before, all symmetric matrices here sum to zero. Their sum is of the form RB_m, because all terms have B_m as their rightmost factor. Since $RB_m = 0$ and B_m is regular, we deduce that $R = 0$, hence we can assume that the new sum contains no symmetric matrices. If m is different from all of i, j, k then $B_i B_j B_k B_m$ is symmetric, hence choosing the right m we deduce that $c_{i,j,k} = 0$. This works unless $n = 4$. If m is equal to one of i, j, k, ℓ then $B_i B_j B_k B_\ell B_m$ equals (up to sign) a product of three different matrices (since the equal matrices cancel), i.e. it is symmetric. Choosing $m = i$ we see that $c_{i,j,k,\ell} = 0$. This always works. Continuing this way, we can eliminate most factors, with the possible exception of the last one, if $n - 1 \equiv 3 \pmod{m}$, i.e. if $4 \mid m$. In this case, we may get $I = cB_1 B_2 \cdots B_{n-1}$. Squaring this equation, remembering that this term is symmetric and cancelling, we get that $I = c^2$, and so $c = \pm 1$.

Let us see where we're standing. Take any linear dependency, containing some term $B_{i_1} B_{i_2} \cdots B_{i_m}$. Multiply the linear dependency by the transpose $B'_{i_m} B'_{i_{m-1}} \cdots B'_{i_1}$ from the right and note that each new term reduces (up to sign) to one of the original 2^{n-1} products (using $B_i B_j = -B_j B_i$ to reorder factors and $B_i^2 = -I$ to cancel identical factors). Hence if $4 \nmid n$ then all matrices are linearly independent, and otherwise, each linear dependency containing $B_{i_1} B_{i_2} \cdots B_{i_m}$ is equivalent to

$$B_{i_1} B_{i_2} \cdots B_{i_m} = cB_1 B_2 \cdots B_{n-1} B_{i_1} B_{i_2} \cdots B_{i_m} = \pm B_{j_1} B_{j_2} \cdots B_{j_{n-1-m}},$$

where the indexes i_k and j_k together form the set $\{1, 2, \ldots, n-1\}$. This inevitably works also the other way around, and so we get a matching of the 2^{n-1} matrices into pairs that are either equal or additive inverses. Note that since $4 \mid n$, either both members are symmetric, or both are skew-symmetric. Let us choose one matrix from each such pair. We claim that the resulting 2^{n-2} matrices are linearly independent.

To prove this, consider first linear dependencies containing I (we assume I was chosen). Since we do not allow I's 'mate' $B_1 B_2 \cdots B_{n-1}$, we can cancel *all* the terms which supposedly equal to I, hence no linear dependency can contain I. Now take any linear dependency. It is easy to see that multiplying any pair by any matrix, we get another pair (the resulting pair index sets are obtained as the symmetric difference with the index set of the matrix, and so still sum to $\{1, 2, \ldots, n-1\}$). Hence multiplying our candidate dependency by the transpose of any of its matrices reduces it to a linear dependency containing I. Since the candidate dependency contained at most one matrix from each pair, all the terms 'survive', and so we reduce to the previous case, proving the claim.

The end is nigh. Since the dimension of the vector space of $n \times n$ matrices over the reals is n^2, we get the inequality $2^{n-2} \leq n^2$ (this can be sharpened to $2^{n-1} \leq n^2$ if $4 \nmid n$). The function 2^{n-2} grows much faster that n^2 and already $2^8 > 10^2$. Hence $n < 10$, and we are left with $n = 2, 4, 6, 8$ (recall we disposed of the case $n = 1$ earlier), since n must be even. We know that all of $n = 2, 4, 8$ are actually possible, so consider the case $n = 6$. Since $4 \nmid 6$, we know that all $2^5 = 32$ products must be linearly independent. How many of them are skew-symmetric? The number is $\binom{5}{1} + \binom{5}{2} + \binom{5}{5} = 5 + 10 + 1 = 16$, exactly half, and they are also linearly independent. Now what is the dimension of the vector space of $n \times n$ skew-symmetric matrices over the reals? Any skew-symmetric real matrix must have all its diagonal elements zero. Therefore, a basis of the skew-symmetric matrices over the reals is $\{O_{ij} \mid 1 \leq i < j \leq n\}$, where O_{ij} has only two non-zero entries, the (i, j)th entry which equals 1, and the (j, i)th entry which equals -1. This basis contains $(n-1) + (n-2) + \cdots + 1 = n(n-1)/2$ elements ($n - 1$ possibilities of j when $i = 1$, and so on up to $i = n - 1$ when j must equal n). In our case, $n = 6$, this basis contains $6 \cdot 5/2 = 15$, hence the 16 skew-symmetric matrices considered earlier can't be linearly independent. Thus the case $n = 6$ is excluded, and the proof is complete.