ON THE GRAM MATRIX

in memory of Maurice Audin

Israel Halperin

(received April 30, 1962)

1. Introduction. The material sketched here is mostly well known and concerns the geometrical inter-relations of vectors in a Hilbert (that is, complete inner product) space. The discussion and references are (obviously) not exhaustive but I hope the reader will find here some interesting problems.

We shall permit the scalars to be real, complex or quaternionic numbers.[1)] The inner product of vectors x and y will be denoted $(x|y)$.

Vectors (sometimes represented by rows of scalars) are to be multiplied by scalars on the left and by linear operators or matrices on the right; thus $(cx|y) = c(x|y)$, $(x|cy) = (x|y)\overline{c}$, $(xA|y) = (x|yA^*)$. If c is a central scalar (that is, $cd = dc$ for all scalars d) we interpret xc to mean cx.

$[x_1, \dots, x_m]$ will denote the subspace spanned by the vectors $x_1, \dots, x_m$.

2. The Gram matrix. An $m \times m$ matrix of scalars $B = (b_{i,k};\ i, k=1, \dots, m)$ will be called a Gram matrix if in some Hilbert space there are vectors $x_1, \dots, x_m$ such that $b_{i,k} = (x_i|x_k)$ for all i, k; then B will be denoted $G(x_1, \dots, x_m)$.

1) For an early memoir using quaternion scalars, see [6]. For more general scalars, see [1].

Canad. Math. Bull. vol. 5, no. 3, September 1962.

As for uniqueness of the $x_1, \ldots, x_m$, it is easy to show that $G(x_1, \ldots, x_m) = G(y_1, \ldots, y_m)$ if and only if there is an isometric mapping of $[x_1, \ldots, x_m]$ onto $[y_1, \ldots, y_m]$ which maps x_i onto y_i for each i.

3. Gram matrices coincide with matrices AA^*. Suppose given a Gram matrix $G(x_1, \ldots, x_m)$. By the Gram-Schmidt orthonormalization procedure[2)] there exist orthonormal vectors $\varphi_1, \ldots, \varphi_n$ with $n \leq m$ such that $[\varphi_1, \ldots, \varphi_n] = [x_1, \ldots, x_m]$. Then $G = AA^*$ where A is the $m \times m$ matrix $(a_{i,k})$ with $a_{i,k} = (x_i | \varphi_k)$ for $k = 1, \ldots, n$, and $a_{i,k} = 0$ for $n < k \leq m$. Actually the Gram-Schmidt procedure shows that the φ_i can be chosen so that A is semi-diagonal, that is: $a_{i,k} = 0$ for all $i < k$.

The converse is trivial: a product AA^* is clearly the Gram matrix of $x_1, \ldots, x_m$ vectors which can be represented (with respect to any orthonormal set of vectors) by the rows of A.

4. Gram matrices coincide with Hermitian definite matrices. We call a matrix $B = (b_{i,k})$ Hermitian if $b_{i,k} = \bar{b}_{k,i}$ for all i, k, definite (abbreviation for positive semi-definite) if $\Sigma_{i,k=1}^{m} t^i b_{i,k} \overline{t^k}$ is real and ≥ 0 for all scalars $t^1, \ldots, t^m$.[3)]

That $G(x_1, \ldots, x_m)$ is Hermitian definite follows from its representation AA^*, or alternatively by direct verification:

$$(x_i | x_k) = \overline{(x_k | x_i)},$$

$$\Sigma_{i,k=1}^{m} t^i (x_i | x_k) \overline{t^k} = ||\Sigma_{i=1}^{m} t^i x_i||^2 \geq 0 .$$

---

2) See [2, pages 30-31], [4, page 16], or [5, page 13].

3) If B is definite then (i): all $b_{i,i}$ are necessarily real and ≥ 0, and (ii): for complex or quaternionic scalars (but not for real scalars) B is necessarily Hermitian.

https://doi.org/10.4153/CMB-1962-027-1 Published online by Cambridge University Press

On the other hand, if B is a given Hermitian definite matrix, then spectral theory shows that $B = A^2$ for some Hermitian definite A and this implies that B is a Gram matrix. But the fact that B is a Gram matrix can be shown without spectral theory. The calculations are cumbersome but nevertheless of some interest; the result is almost trivial when $m = 1$ and follows by induction from the following lemma.

(1) Suppose $x_1, \ldots, x_m$ are given vectors and $B = (b_{i,k})$ is a Hermitian definite $(m+1) \times (m+1)$ matrix with $b_{i,k} = (x_i | x_k)$ for $i, k = 1, \ldots, m$. Then there exists a vector x_{m+1} such that $b_{i,k} = (x_i | x_k)$ for $i, k = 1, \ldots, m+1$.

Every vector has the form: $d^1 x_1 + \ldots + d^m x_m + y$ (with suitable scalars $d^1, \ldots, d^m$ and a suitable vector y orthogonal to all of $x_1, \ldots, x_m$) and (1) can be deduced from the following lemma:

(2) Suppose $B = (b_{i,k})$ is a Hermitian definite $(m+1) \times (m+1)$ matrix. Then there exist scalars $d^1, \ldots, d^m$ such that

$d^1 b_{1,j} + \ldots + d^m b_{m,j} = b_{m+1,j}$ for $j = 1, \ldots, m$;

$d^1 b_{1,m+1} + \ldots + d^m b_{m,m+1}$ is real and $\leq b_{m+1,m+1}$;

then in (1), x_{m+1} may be taken as $d^1 x_1 + \ldots + d^m x_m + y$ where y is any vector orthogonal to all of $x_1, \ldots, x_m$ which satisfies $||y||^2 = b_{m+1,m+1} - ||d^1 x_1 + \ldots + d^m x_m||^2$

$(= b_{m+1,m+1} - (d^1 b_{1,m+1} + \ldots + d^m b_{m,m+1}))$.

To prove (2), choose n to be the largest integer with the property: for some set of integers $i_1, \ldots, i_n$ with $1 \leq i_1 < i_2 < \ldots < i_n < m$ the row-vectors

$$(b_{i_1,i_1}, \ldots, b_{i_1,i_n}), \ldots, (b_{i_n,i_1}, \ldots, b_{i_n,i_n})$$

are linearly independent. Necessarily $0 \leq n \leq m$.

https://doi.org/10.4153/CMB-1962-027-1 Published online by Cambridge University Press

We note that if for some $i \leq m+1$, $b_{i,i} = 0$, then $b_{i,k} = 0 = b_{k,i}$ for all $k = 1, \ldots, m+1$; to see this, observe that if $k \neq i$, then $sb_{i,i}\bar{s} + sb_{i,k}\bar{t} + tb_{k,i}\bar{s} + tb_{k,k}\bar{t} \geq 0$ for all scalars s, t. Now choose $t = -\epsilon$ with ϵ real and > 0 and choose $s = b_{k,i}$. Then since $b_{i,i} = 0$ it follows that $\epsilon^2 b_{k,k} \geq 2\epsilon |b_{i,k}|^2$; hence $\epsilon b_{k,k} \geq 2|b_{i,k}|^2$ for all $\epsilon > 0$ and hence also for $\epsilon = 0$. This proves: $b_{i,k} = 0$ as stated.

Now, if $n = 0$ then $b_{i,i} = 0$ for all $i \leq m$, hence $b_{i,k} = 0$ for all $i, k = 1, \ldots, m+1$, except possibly for $b_{m+1,m+1}$. In this case (2) clearly holds with $d^i = 0$ for $i = 1, \ldots, m$.

Now suppose $n > 0$. We may suppose $i_r = r$ for $1 \leq r \leq n$. Then, for suitable scalars $c^1, \ldots, c^n$ we must have:

(3) $\quad b_{m+1,k} = c^1 b_{1,k} + \ldots + c^n b_{n,k}$ <u>for</u> $k = 1, \ldots, n$.

We shall show:

(4) $\quad c^1 b_{1,m+1} + \ldots + c^n b_{n,m+1}$ <u>is real and</u> $\leq b_{m+1,m+1}$;

(5) <u>For all</u> $n+1 \leq k \leq m$, (3) <u>holds.</u>

This will imply that (2) holds with $d^i = c^i$ for $1 \leq i \leq n$ and $d^i = 0$ for $n+1 \leq i \leq m$.

To prove (4), note that for every scalar t,

$$(6) \quad \sum_{i,k=1}^{n} c^i b_{i,k} \overline{c^k} + \sum_{k=1}^{n} t b_{m+1,k} \overline{c^k} + \sum_{k=1}^{n} b_{k,m+1} \bar{t} + t b_{m+1,m+1} \bar{t}$$

is real and ≥ 0.

https://doi.org/10.4153/CMB-1962-027-1 Published online by Cambridge University Press

By (3), $\sum_{k=1}^{n} b_{m+1,k}\overline{c^k} = \sum_{i,k=1}^{n} c^i b_{i,k}\overline{c^k}$; hence (6), with $t = 1$, yields (4).

To prove (5) suppose j fixed with $n < j \leq m$. Then the row vector $(b_{j,1}, \ldots, b_{j,n}, b_{j,j})$ must be a left linear combination of $(b_{i,1}, \ldots, b_{i,n}, b_{i,j})$, $i = 1, \ldots, n$, because of the maximal property of n . Thus for suitable scalars $d^1, \ldots, d^n$:

(7) $\quad b_{j,k} = d^1 b_{1,i} + \ldots + d^n b_{n,i}$ <u>for</u> $i = 1, \ldots, n, j$.

We shall first show that (7) holds for $i = m + 1$ also. Since B is definite,

$$\sum_{i,k=1,\ldots,n,j,m+1} t^i b_{i,k}\overline{t^k} \geq 0$$

for all scalars $t^1, \ldots, t^n, t^j, t^{m+1}$. Choose $t^i = d^i$ for $1 \leq i \leq n$, $t^j = -1$ and $t^{m+1} = \epsilon t$ with ϵ real, > 0 . Then since $b_{j,j} = \sum_{i=1}^{n} d^i b_{i,j} = \sum_{i,k=1}^{n} d^i b_{i,k}\overline{d^k}$, therefore

$$b_{j,j} - b_{j,j} - b_{j,j} + b_{j,j} + \epsilon \sum_{i=1}^{n} d^i b_{i,m+1}\overline{t} + \epsilon \sum_{i=1}^{n} t b_{m+1,i}\overline{d^i}$$

$$- \epsilon\, b_{j,m+1}\overline{t} - \epsilon t b_{m+1,j} + \epsilon^2 |t|^2 b_{m+1,m+1} \geq 0$$

for every scalar t and all $\epsilon > 0$. Hence

$$(b_{j,m+1} - \sum_{i=1}^{n} d^i b_{i,m+1})\overline{t} + t(b_{m+1,j} - \sum_{i=1}^{n} b_{m+1,i}\overline{d^i}) \leq \epsilon |t|^2 b_{m+1,m+1}$$

for all $\epsilon > 0$, and hence also for $\epsilon = 0$, and for every scalar t.

Now with $t = b_{j,m+1} - \sum_{i=1}^{n} d^i b_{i,m+1}$ we obtain:

https://doi.org/10.4153/CMB-1962-027-1 Published online by Cambridge University Press

$2|b_{j,m+1} - \sum_{i=1}^{n} d^i b_{i,m+1}| = 0$ and hence (7) holds for $i = m+1$.

Now returning to (5), we have:

$$b_{m+1,j} = \sum_{k=1}^{n} b_{m+1,k} \overline{d^k} = \sum_{i,k=1}^{n} c^i b_{i,k} \overline{d^k} = \sum_{i=1}^{n} c^i b_{i,j}$$

so (3) holds for $k = j$, as required.

This completes the proof of (2), and (1), and shows that B is a Gram matrix if and only if B is Hermitian definite.

5. <u>Volume of a parallelepiped</u>. If $x_1, \ldots, x_m$ are given vectors we define the parallelepiped $P(x_1, \ldots, x_m)$ to consist of all vectors of the form $\sum_{i=1}^{m} c^i x_i$ with c^i real and $0 \leq c^i \leq 1$ for all i.

The high-school formula for m-dimensional volume of a parallelepiped (we shall use the symbol $V_m(x_1, \ldots, x_m)$) would be defined (presumably) by induction on m as follows:

(8) $\quad V_1(x_1) = ||x_1||$,

$$V_{m+1}(x_1, \ldots, x_{m+1}) = ||x'_{m+1}|| V_m(x_1, \ldots, x_m)$$

where x'_{m+1} is the component of x_{m+1} orthogonal to $[x_1, \ldots, x_m]$.

The formulae (8) determine $V_m(x_1, \ldots, x_m)$ uniquely as a non-negative real number which (clearly) is 0 if and only if the vectors $x_1, \ldots, x_m$ are linearly dependent. It is usually assumed that the value of $V_m(x_1, \ldots, x_m)$ does not depend on the order of $x_1, \ldots, x_m$ and this requires proof.

An easy proof is possible if determinants are available, that is, if the scalars are real or complex numbers. Assuming

https://doi.org/10.4153/CMB-1962-027-1 Published online by Cambridge University Press

the vectors $x_1, \ldots, x_m$ are linearly independent, we can use the Gram-Schmidt orthonormalization to obtain

(9) $\quad x_i = a_{i,1}\varphi_1 + \ldots + a_{i,i}\varphi_i$ for $i = 1, \ldots, m$

with $a_{i,i}$ real and > 0 for all i. Then $x_i' = a_{i,i}\varphi_i$, $||x_i'|| = a_{i,i}$, and $V_m(x_1, \ldots, x_m) = \pi_{i=1}^{m} a_{i,i} = \det A$ where A is the $m \times m$ matrix which has i,k-th entry equal to $a_{i,k}$ if $1 \leq k \leq i$ and has all other entries equal to 0. Now we have the equality:

$$V_m(x_1, \ldots, x_m) = \sqrt{\det G(x_1, \ldots, x_m)}$$

(which is also obviously valid if the vectors $x_1, \ldots, x_m$ are linearly dependent); this implies that the value of V_m is independent of the order of $x_1, \ldots, x_m$, provided the scalars are real or complex numbers.

The following direct proof is valid for real, complex or quaternionic scalars. Clearly we need only show that if $m > 1$ then the value of V_m is unchanged when x_m and x_{m-1} are permuted. We will then have two versions of (9) to consider, say:

$$x_{m-1} = a_{m-1,1}\varphi_1 + \ldots + a_{m-1,m-2}\varphi_{m-2} + a_{m-1,m-1}\varphi_{m-1},$$

$$x_m = a_{m,1}\varphi_1 + \ldots + a_{m,m-2}\varphi_{m-2} + a_{m,m-1}\varphi_{m-1} + a_{m,m}\varphi_m;$$

and

$$x_m = a_{m,1}\varphi_1 + \ldots + a_{m,m-2}\varphi_{m-2} + a'_{m,m-1}\psi_{m-1},$$

$$x_{m-1} = a_{m-1,1}\varphi_1 + \ldots + a_{m-1,m-2}\varphi_{m-2} + a'_{m-1,m-1}\psi_{m-1}$$
$$+ a'_{m-1,m}\psi_m;$$

and we need only show that $a_{m-1,m-1}a_{m,m} = a'_{m,m-1}a'_{m-1,m}$.

https://doi.org/10.4153/CMB-1962-027-1 Published online by Cambridge University Press

We have:

$$a_{m-1,m-1}(\varphi_{m-1}|\psi_{m-1}) = a'_{m-1,m-1} ,$$

$$a'_{m,m-1}(\psi_{m-1}|\varphi_{m-1}) = a_{m,m-1} ,$$

$$|a_{m-1,m-1}|^2 = |a'_{m-1,m-1}|^2 + |a'_{m-1,m}|^2 ,$$

$$|a'_{m,m-1}|^2 = |a_{m,m-1}|^2 + |a_{m,m}|^2 ;$$

so

$$a_{m-1,m-1}|a_{m,m-1}| = a'_{m,m-1}|a'_{m-1,m-1}| ,$$

$$(a_{m-1,m-1})^2((a'_{m,m-1})^2 - (a_{m,m})^2)$$

$$= (a'_{m,m-1})^2((a_{m-1,m-1})^2 - (a'_{m-1,m})^2) ,$$

$$(a_{m-1,m-1})^2(a_{m,m})^2 = (a'_{m,m-1})^2(a'_{m-1,m})^2$$

and hence $a_{m-1,m-1}\,a_{m,m} = a'_{m,m-1}\,a'_{m-1,m}$ as required.

6. Determinants of matrices with non-commutative scalars. There are well known obstacles to an extension of the usual theory of determinants for matrices to the case where the entries are taken from a non-commutative ring.[4] But if A is an $m \times m$ matrix of quaternions we can define a generalization of "the absolute value of the determinant," denoted $D(A)$, as follows. We observe that

(10) the quotient $\dfrac{V_m(x_1A, \ldots, x_mA)}{V_m(x_1, \ldots, x_m)}$ is a constant if $x_1, \ldots, x_m$ are arbitrary linearly independent row vectors (each of m scalars).

Then we define $D(A)$ to be the value of the quotient in (10).

---

4) See [3] and the references given there.

The reader can use the identities:

$V_m(c_1x_1, \ldots, c_mx_m) = |c_1 \ldots c_m| V_m(x_1, \ldots, x_m)$;

$V(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_m) = V(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_m)$

<u>whenever</u> $y = x^i$ + <u>an arbitrary left linear combination of</u> $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_m$

to prove (10) by showing that given linearly independent vectors $x_1, \ldots, x_m$ can be transformed into given linearly independent vectors $y_1, \ldots, y_m$ by successive steps each of which leaves the quotient in (10) unaltered.

The reader can also verify that the function $D(A)$ has the properties:

(11) $D(A)$ is defined and is a real number ≥ 0, $D(0) = 0$, $D(1) = 1$;

(12) If A is a semi-diagonal matrix $(a_{i,k})$ (that is, $a_{i,k} = 0$ for all $i < k$) then $D(A) = \pi_{i=1}^{m} |a_{i,i}|$;

(13) $D(A) = 0$ if and only if $xA = 0$ for some non-zero vector x;

(14) $D(AB) = D(A)D(B)$;[5)]

(15) $D(AA^*) = D(A^*A)$[6)], $D(A) = D(A^*) = \sqrt{D(AA^*)}$;

---

5) The validity of (14) follows from (10) if $D(A) \neq 0$ and from (13) if $D(A) = 0$.

6) Use: $A^*A = UAA^*U^*$ for some unitary U. Alternatively: there are independent vectors $x_1, \ldots, x_m$ such that for each i, $x_iAA^* = \lambda_i x_i$ for some real $\lambda_i \geq 0$ (the x_i can even be chosen to be orthonormal); then $y_iA^*A = \lambda_i y_i$ if y_i is chosen to be x_iA. Hence $D(A^*A) = \pi_i \lambda_i = D(AA^*)$ if the y_i are linearly independent (and also, by an obvious argument, if the y_i are linearly dependent).

(16) $D(A)$ = square root of the product of the spectral values of AA^*.

We again have $V(x_1, \ldots, x_m) = \sqrt{D(G(x_1, \ldots, x_m))}$.

7. Generalized Schwarz inequalities. Suppose now that B is Hermitian. We shall say that B'' is obtained from B by a symmetrical transformation if some B' is obtained by adding to the i-th row of B a left linear combination $\sum_{j \neq i} c^j(b_{j,k};\ k=1, \ldots, m)$ of the other rows of B and then B'' is obtained by adding to the i-th column of B' the corresponding conjugate right linear combination $\Sigma(b'_{k,j};\ k=1, \ldots, m)\overline{c^j}$ of the other columns of B' (symmetrical transformation of the matrix B corresponds to a certain change in the basis of the vector space).

If B_o is an $m \times m$ diagonal matrix (that is, $(b_o)_{i,k} = 0$ if $i \neq k$), we shall say that B_o is diagonally related to B if B_o can be obtained from B by a succession of symmetrical transformations. It is easy to see that every Hermitian B is diagonally related to some diagonal B_o (necessarily Hermitian but in general not uniquely determined by B) and when B is Hermitian definite then B_o is Hermitian definite and the product of the diagonal entries of B_o is non-negative, coinciding with $D(B)$.

We are led to define by induction on m for a given Hermitian $m \times m$ matrix B an expression $E(B)$ which is a certain polynomial in the entries $b_{i,k}$, $i, k=1, \ldots, m$, as follows:

If $m = 1$ and $B = (b_{1,1})$, set $E(B) = b_{1,1}$.

If $m > 1$ and $E(B')$ is already defined for every $(m-1) \times (m-1)$ Hermitian matrix B', set $E(B) = E(B')$ where B' is the $(m-1) \times (m-1)$ Hermitian matrix with $(b')_{i,k} = b_{1,1} b_{i,k} - b_{i,1} b_{1,k}$ for $2 \leq i, k \leq m$.

https://doi.org/10.4153/CMB-1962-027-1 Published online by Cambridge University Press

It is not difficult to prove that B is definite if and only if $E(B') \geq 0$ for every submatrix $B' = (b'_{i,k};\ i, k = i_1, \ldots, i_n)$ with $1 \leq i_1 < \ldots < i_n \leq m$.

If $m = 2$, then $E(B)$ is simply $b_{1,1}b_{2,2} - |b_{1,2}|^2$, which corresponds to the well known Schwarz inequality: $(x_1|x_1)(x_2|x_2) - |(x_1|x_2)|^2 \geq 0$. We may therefore call each inequality: $E(B) \geq 0$ a generalized Schwarz inequality.

8. The Hermitian matrix as an operator. Suppose B is an $m \times m$ Hermitian matrix and for each row vector x let xB be the row obtained by matrix multiplication. Then B determines a self-adjoint operator on the m dimensional Hilbert space of row vectors.

Elementary spectral theory shows that B is a difference of two Hermitian definite operators $B = G(x_1, \ldots, x_m) - G(y_1, \ldots, y_m)$, say, such that $G(x_1, \ldots, x_m)\, G(y_1, \ldots, y_m) = 0$.

It is interesting to enquire into the behaviour of $x_1, \ldots, x_m,\ y_1, \ldots, y_m$ when B is replaced by $B - tI$ with t a real number and I the $m \times m$ unit matrix (in particular, when B is definite). Also, what is the behaviour of the spectral vectors and spectral values when B undergoes a symmetric transformation? If $G(x_1, \ldots, x_m)$ and $G(y_1, \ldots, y_m)$ commute then their product matrix is of the form $G(z_1, \ldots, z_m)$; can the z_i be expressed in a simple way in terms of $x_1, \ldots, x_m$, $y_1, \ldots, y_m$?

9. Special Gram Matrices. It is interesting to investigate families of vectors $x_1, \ldots, x_m$ for which

(17) $(x_i|x_k) = d$ (a constant, necessarily real) for all $i \neq k$.

If $d = 0$, this simply means that the vectors are mutually

https://doi.org/10.4153/CMB-1962-027-1 Published online by Cambridge University Press

orthogonal. Then, as is well known, such vectors $x_1, \ldots, x_m$ do exist, with arbitrary non-negative values for $(x_i|x_i)$; in this case, the x_i (if non-zero) are necessarily linearly independent and, if the Hilbert space is of dimension $> m$, it is possible to adjoin a vector x_{m+1} so that $x_1, \ldots, x_{m+1}$ also have property (17).

Suppose $d < 0$, that is, $d = -c$ with $c > 0$. Then vectors do exist satisfying (17) but the values of $(x_i|x_i)$ cannot be assigned without restriction. They are arbitrary, subject to the following condition:

$$(18) \quad \sum_{i=1}^{m} \frac{c}{c + ||x_i||^2} \leq 1 \ .$$

Moreover, if equality holds in (18) then the vectors are linearly dependent and it is impossible to adjoin a vector x_{m+1} so that (17) continues to hold. But if strict inequality holds in (18) then the vectors are linearly independent and (if the dimension of the Hilbert space $> m$) it is possible to adjoin x_{m+1} so that (17) continues to hold.

To show this we may assume $c = 1$ (by replacing the former x_i by $\sqrt{c}x_i$). Suppose now that $x_1, \ldots, x_m$ satisfy (17); then any vector (x_{m+1}, say) can be expressed as $g + y$ with $g = a_1 x_1 + \ldots + a_m x_m$ and some y orthogonal to each of $x_1, \ldots, x_m$.

The statement $(x_{m+1}|x_i) = (g|x_i) = -1$ for all $i = 1, \ldots, m$ is equivalent, by (17), to $\sum_{j=1}^{m} a_j(x_j|x_i) = -1$, that is, to:

$$(19) \quad a_i(||x_i||^2 + 1) = (\sum_{j=1}^{m} a_j) - 1 \text{ for } i=1, \ldots, m \ .$$

Elementary manipulation of fractions show that (19) is equivalent to:

$$(20) \quad \sum_{j=1}^{m} \frac{1}{1+||x_j||^2} < 1 \quad \underline{\text{and}}$$

$$a_i = \frac{-1}{||x_i||^2+1} \left(1 - \sum_{j=1}^{m} \frac{1}{1+||x_j||^2}\right)^{-1}$$

However (20) may also be obtained from the calculations:

$$||g||^2 = (g|g) = \sum_{j=1}^{m} (a_j x_j | g) = \sum_{j=1}^{m} a_j(-1) = -\sum_{j=1}^{m} a_j \ ;$$

$$-a_i = \frac{1+||g||^2}{1+||x_i||^2} \ ;$$

$$||g||^2 = (1+||g||^2) \sum_{j=1}^{m} \frac{1}{1+||x_j||^2} \ ;$$

$$\sum_{i=1}^{m} \frac{1}{1+||x_i||^2} + \frac{1}{1+||g||^2} = 1 \ .$$

Since $||x_{m+1}||^2 = ||g||^2 + ||y||^2$ it is clear that if x_{m+1} exists as described, then $\sum_{i=1}^{m+1} \frac{1}{1+||x_i||^2} \leq 1$ with equality equivalent to linear dependence of x_{m+1} on $x_1, \ldots, x_m$. On the other hand, (20) shows that if $\sum_{j=1}^{m} \frac{1}{1+||x_j||^2} < 1$ then x_{m+1} does exist and, if the dimension of the Hilbert space is greater than m, the value of $||x_{m+1}||$ is arbitrary, subject to the condition (18) (for $m+1$).

https://doi.org/10.4153/CMB-1962-027-1 Published online by Cambridge University Press

As a corollary, we see that a sequence of vectors x_n with $(x_n | x_m) = -1$ for all $n \neq m$ can exist in a Hilbert space if and only if the dimension of the space is infinite. Then the vectors are necessarily linearly independent and the values of $||x_n||$ are arbitrary except for the condition:

$$(21) \quad \sum_{n=1}^{\infty} \frac{1}{1 + ||x_n||^2} \leq 1 .$$

The necessity of the condition (21) was first observed by I. Amemiya (oral communication, 1959). His elegant proof ran as follows: choose a unit vector g orthogonal to all the given x_n (enlarge the Hilbert space if necessary). Then the vectors $\left(\frac{1}{1 + ||x_n||^2}\right)^{1/2} (g + x_n)$, $n = 1, 2, \ldots$ are orthonormal and Bessel's inequality, applied to g with respect to this orthonormal set, gives (21).

The writer then used the Gram-Schmidt orthonormalization process to prove that for given real $p_{n,m}$ with all $p_{n,n} > 0$ and $p_{n,m} \leq -1$ for all $n \neq m$, the condition $\sum_{n=1}^{\infty} \frac{1}{1 + p_{n,n}} \leq 1$ is necessary and sufficient for the existence of vectors x_n with $(x_n | x_m) = p_{n,m}$ for all n, m (later, A. Renyi found a proof of Amemiya's result using Bessel's inequality directly).

Then Amemiya remarked (in a letter dated October, 1961) that the following variation of his method also showed that (21) was a sufficient condition if $p_{n,m} = -1$ for all $n \neq m$:

For given $d_n > 0$ $(n \geq 1)$ with $\sum_{n=1}^{\infty} \frac{1}{1 + d_n^2} \leq 1$ simply choose

https://doi.org/10.4153/CMB-1962-027-1 Published online by Cambridge University Press

y_n, $n \geq 0$ to be orthogonal vectors with $||y_o|| = 1$, $||y_n||^2 = (1 + d_n^2)^{-1}$, let $g = ay_o - \sum_{n=1}^{\infty} y_n$ with $a \geq 0$, $a^2 = 1 - \sum_{n=1}^{\infty} \frac{1}{1 + d_n^2}$. Then $x_n = g + (1 + d_n^2)y_n$, $n \geq 1$ satisfies $||x_n|| = d_n$, $(x_n|x_m) = -1$ for $n \neq m$.

The reader may find it interesting to vary this last construction of Amemiya in order to deal with arbitrary real $p_{n,m} \leq -1$.

The reader may also seek to prove, by direct arithmetic, the following consequence of Amemiya's theorem: if $p_n > 0$ for $n = 1, 2, \ldots$ and $q_1 = p_1$, $q_{n+1} = p_{n+1}(1 + q_1^2 + \ldots + q_n^2)$ for $n \geq 1$, then $\sum_{n=1}^{\infty} \frac{p_n^2}{1 + q_n^2} \leq 1$.

REFERENCES

1. N. Bourbaki, Algèbre, Ch. 9, §7.

2. R. Courant and D. Hilbert, Methoden der Mathematischen Physik, vol. 1 (Springer, Berlin, 1931).

3. J. Dieudonné, Les déterminants sur un corps non commutatif, Bull. Soc. Math. France, vol. 71 (1943), pp. 27-45.

4. John von Neumann, Functional Operators, vol. II: The Geometry of Orthogonal Spaces, Annals of Mathematics Studies, No. 22, Princeton, 1950.

5. M.H. Stone, Linear Transformations in Hilbert Space, American Mathematical Society Colloquium Publications Vol. XV, New York, 1932.

6. O. Teichmuller, Operatoren im Wachsschen Raum, J. Reine und Angew. Math., vol. 174 (1935), pp. 73-124.

Queen's University
Kingston, Canada