Can. J. Math., Vol. XL, No. 1, 1988, pp. 38-54

N SUBSPACES

V. S. SUNDER

Introduction. It is a well-known fact (cf., for instance Lemma 7.3.1 of [**8**], and also [**2**] and [**4**]) that if M and N are closed subspaces of a finite-dimensional Hilbert space, and if M and N are in 'generic' position (i.e., any two of the four subspaces M, $M^{\perp}$, N, $N^{\perp}$ have trivial intersection), then N is the graph of a linear isomorphism of M onto $M^{\perp}$. To be sure, there exist infinite-dimensional versions of this, where one must allow for unbounded operators in case the 'gap' between M and N is zero, in the sense of Kato [**7**]. (There is an extensive literature on pairs of subspaces, [**2**], [**3**], [**4**], [**6**] and [**7**], to cite a few; for a fairly extensive bibliography, see [**3**].)

This paper addresses itself to the case of n $(2 \leqq n < \infty)$ subspaces. Theorem 1 generalises the assertion of the preceding paragraph as follows: if $M_1, \ldots, M_n$ are closed subspaces of a Hilbert space H such that H is the algebraic direct sum of the M_i's, then there exists an orthogonal direct sum decomposition

$$H = L_1 \oplus \ldots \oplus L_n$$

such that M_k looks like the graph of a bounded linear transformation from L_k into $L_1 \oplus \ldots \oplus L_{k-1}$ for $1 \leqq k \leqq n$.

The orthogonal projection onto M_k is explicitly computed in terms of the above operator, and this description is used to attack the problem of unitary equivalence for n-tuples of closed subspaces. In a certain 'generic' case (see Definition 1), the above problem reduces to the unitary equivalence problem for single operators. As a by-product of the above computations, one has a concrete description of the commutant $\{P_1, \ldots, P_n\}'$ (where P_i = projection on M_i), which leads easily to examples of sets $\{P_1, \ldots, P_n\}$ of n projections, with $n \geqq 3$, such that $\mathscr{B}(H)$ is generated as a von Neumann algebra by $\{P_1, \ldots, P_n\}$ but by no proper subset. (For a specific example with $n = 3$, see [**1**].)

The final section of the paper applies the machinery developed earlier to solve the statistical problem of computing the canonical partial correlation coefficients between three sets of random variables (cf. [**9**]).

Acknowledgement. Some of the results, as well as the title, of this paper are inspired by (and constitute natural generalisations of) [**4**]. The author

---

Received April 9, 1985 and in revised form July 2, 1985.

would like to acknowledge his gratitude to S. K. Mitra for several stimulating conversations, particularly in connection with the statistical application in the last section of the paper. The author would also like to thank the referee for some helpful comments as well as for bringing some of the references to his awareness.

Notation. Throughout this paper, the symbols $M_1, \dots, M_n$ will denote closed subspaces of a (real or complex) Hilbert space H such that

$$H = M_1 + \dots + M_n$$

and (*)

$$M_i \cap \sum_{j \neq i} M_j = \{0\}.$$

(As a matter of convention, we shall employ the symbol

$$H = \bigoplus_{i=1}^{n} L_i$$

only when the subspaces L_i are mutually orthogonal and together span. To distinguish from such an orthogonal direct sum, we shall say that

$$H = \sum_{i=1}^{n} M_i$$

is an algebraic direct sum if the closed subspaces $M_1, \dots, M_n$ satisfy condition (*)). For $1 \leqq k \leqq n$, define

$$S_k = \sum_{j=1}^{k} M_j, \quad L_k = S_k \cap S_{k-1}^{\perp}$$

with the understanding that $S_0 = \{0\}$, so that $L_1 = S_1$. It is clear that the L_k's are pairwise orthogonal subspaces of H such that

$$\sum_{j=1}^{k} M_j = \bigoplus_{j=1}^{k} L_j;$$

in particular,

$$H = \bigoplus_{k=1}^{n} L_k.$$

(The passage from the M_k's to the L_k's may be viewed as a Gram-Schmidt orthogonalisation process for subspaces.)

The orthogonal projections onto M_k and L_j will be denoted by P_k and E_j respectively. For $1 \leqq j \leqq k \leqq n$, define

$$A_{jk}:M_k \to L_j$$

by

$$A_{jk}x = E_jx.$$

Thus, A_{jk} is just the operator E_j, but viewed as operating between the Hilbert spaces M_k and L_j.

Finally, with respect to the decomposition

$$H = \bigoplus_{j=1}^{n} L_j,$$

let P_k be represented by the operator matrix $P_k = ((C_{k,ij})$, where of course, $C_{k,ij}$ is the unique operator from L_j to L_i satisfying

$$\langle C_{k,ij}x, y\rangle = \langle P_kx, y\rangle \quad \text{for all } x \text{ in } L_j, y \text{ in } L_i.$$

The main result.

LEMMA 1. *Fix $k \leqq n$. Then,*

$$C_{k,ij} = \begin{cases} A_{ik}A_{jk}^* & \text{if } 1 \leqq i, j \leqq k \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Since $M_k \subseteq L_1 \oplus \ldots \oplus L_k$, it is clear that $C_{k,ij} = 0$ if $i > k$ or $j > k$. So, fix $i, j \leqq k$. Note first that $A_{ik}A_{jk}^*$ is an operator from L_j to L_i; for arbitrary $x \in L_j$ and $y \in L_i$, note that

$$\begin{aligned} \langle A_{ik}A_{jk}^*x, y\rangle &= \langle E_iA_{jk}^*x, y\rangle \\ &= \langle A_{jk}^*x, y\rangle \qquad (\text{since } y \in L_i) \\ &= \langle A_{jk}^*x, P_ky\rangle \\ &= \langle x, A_{jk}P_ky\rangle \\ &= \langle x, E_jP_ky\rangle \\ &= \langle x, P_ky\rangle \qquad (\text{since } (x \in L_j) \\ &= \langle P_kx, y\rangle. \end{aligned}$$

LEMMA 2. *For $1 \leqq k \leqq n$, A_{kk} is an invertible operator from M_k to L_k.*

Proof. Since

$$L_k \subseteq \sum_{j=1}^{k} M_j,$$

it follows that E_k maps $\sum_{j=1}^{k} M_j$ onto L_k. However,

$$\sum_{j=1}^{k-1} M_j = \bigoplus_{j=1}^{k-1} L_j \subseteq L_k^{\perp}$$

and hence E_k annihilates $\sum_{j=1}^{k-1} M_j$. It follows that E_k maps M_k onto L_k; i.e., A_{kk} is onto.

Next, if $x \in M_k$ is such that $E_k x = 0$, it follows that

$$x \in M_k \cap L_k^{\perp} = M_k \cap \sum_{j=1}^{k-1} M_j,$$

which contradicts the standing assumption that H is the algebraic direct sum of the M_j's, unless $x = 0$; i.e., A_{kk} is one-to-one.

THEOREM 1. *Let M_i, L_i be as above. Then, there exist bounded operators $B_{jk}:L_k \to L_j$ for $1 \leqq j \leqq k \leqq n$ such that, with respect to the decomposition*

$$H = \bigoplus_{j=1}^{n} L_j,$$

one has, for $1 \leqq k \leqq n$,

(1) $$M_k = \{(B_{1k}x, B_{2k}x, \dots, B_{k-1,k}x, x, 0, \dots, 0):x \in L_k\}.$$

Proof. With the notation already established, define

$$B_{jk} = A_{jk} \circ A_{kk}^{-1}, \quad \text{for } 1 \leqq j \leqq k \leqq n.$$

The boundedness of B_{jk} follows from Lemma 2 and the open mapping theorem. Observe also that, by the definition of the A_{jk}'s and the B_{jk}'s,

$$\begin{aligned} M_k &= \{(A_{1k}x, \dots, A_{k-1,k}x, A_{kk}x, 0, 0, \dots, 0):x \in M_k\} \\ &= \{(B_{1k}x, \dots, B_{k-1,k}x, x, 0, \dots, 0):x \in L_k\}, \end{aligned}$$

again by Lemma 2.

Remark 1. (a) Note that $B_{kk} = I_{L_k}$.

(b) In the converse direction to Theorem 1, note that if

$$H = \bigoplus_{j=1}^{n} L_j$$

is an orthogonal direct sum decomposition of H, if $B_{jk}:L_k \to L_j$ are arbitrary bound operators, for $1 \leqq j < k \leqq n$, and if M_k is defined by (1), then H is the algebraic direct sum of the M_k's and the above process applied to the M_k's will yield the given L_j's and B_{jk}'s.

LEMMA 3. *With respect to the decomposition*

$$H = \bigoplus_{j=1}^{n} L_j,$$

the projection P_k onto M_k is given by the operator matrix $(C_{k,ij})$, where

$$C_{k,ij} = \begin{cases} B_{ik}\left(\sum_{l=1}^{k} B_{lk}^* B_{lk}\right)^{-1} B_{jk}^* & \textit{for } 1 \leqq i, j \leqq k \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. Note first that $B_{kk} = I_{L_k}$ and hence the operator

$$\sum_{l=1}^{k} B_{lk}^* B_{lk}$$

is invertible.

For $j \leqq k$, we have $A_{jk} = B_{jk} \circ A_{kk}$, by definition. So Lemma 1 shows that

$$C_{k,ij} = B_{ik} \circ A_{kk} \circ A_{kk}^* \circ B_{jk}^* \quad \text{for } 1 \leqq i, j \leqq k,$$

and $C_{k,ij} = 0$ if $i > k$ or $j > k$. Hence, to prove the lemma, it suffices to establish that

$$A_{kk} \circ A_{kk}^* = \left(\sum_{l=1}^{k} B_{lk}^* B_{lk}\right)^{-1}.$$

To see this, start from the obvious equality

$$\sum_{j=1}^{k} A_{jk}^* \circ A_{jk} = I_{M_k},$$

and substitute $A_{jk} = B_{jk} \circ A_{kk}$ to conclude

$$I_{M_k} = \sum_{j=1}^{k} A_{kk}^* \circ B_{jk}^* \circ B_{jk} \circ A_{kk} = A_{kk}^*\left(\sum_{j=1}^{k} B_{jk}^* B_{jk}\right)A_{kk},$$

whence

$$A_{kk}^{*-1} \cdot A_{kk}^{-1} = \sum_{j=1}^{k} B_{jk}^* B_{jk},$$

i.e.,

$$(A_{kk} A_{kk}^*)^{-1} = \sum_{j=1}^{k} B_{jk}^* B_{jk},$$

as desired.

https://doi.org/10.4153/CJM-1988-002-0 Published online by Cambridge University Press

2. The unitary equivalence problem.

THEOREM 2. *Let* $M_1, \ldots, M_n$ (resp., $M'_1, \ldots, M'_n$) *be closed subspaces of* H *such that* H *is the algebraic direct sum of the* M_i*'s (resp.,* M'_i*'s). Let*

$$H = \bigoplus_{j=1}^{n} L_j \quad \left(\text{resp., } H = \bigoplus_{j=1}^{n} L'_j\right)$$

be the associated orthogonal decomposition, and let $B_{jk}:L_k \to L_j$ *(resp.,* $B'_{jk}:L'_k \to L'_j$*) be the operators given by Theorem* 1.

(a) *If* U *is a unitary operator on* H *such that* $U(M_i) = M'_i$ *for* $1 \leqq i \leqq n$, *then* $U(L_i) = L'_i$ *for* $1 \leqq i \leqq n$. *If* $U_i:L_i \to L'_i$ *is the restriction of* U *to* L_i, *then*

$$U_i B_{ik} = B'_{ik} U_k \quad \text{for } 1 \leqq i < k \leqq n.$$

(b) *Conversely, if* $U_i:L_i \to L'_i$ *are unitary operators such that*

$$U_i B_{ik} = B'_{ik} U_k \quad \text{for } 1 \leqq i < k \leqq n,$$

then, there exists a unitary operator U *on* H *such that*

$$U|L_i = U_i \text{ and } U(M_i) = M'_i \quad \text{for } 1 \leqq i \leqq n.$$

Proof. (a) If U is a unitary operator on H such that $U(M_i) = M'_i$ for each i, it is easy to see that $U(L_i) = L'_i$ for each i. The hypothesis $U(M_k) = M'_k$ is clearly equivalent to $UP_k = P'_k U$ (where, of course P'_k is the projection onto M'_k). It follows now from Lemma 3, that, for $1 \leqq i, j \leqq k$,

$$(2) \quad U_i B_{ik}\left(\sum_{l=1}^{k} B_{lk}^* B_{lk}\right)^{-1} B_{jk}^* = B'_{ik}\left(\sum_{l=1}^{k} B_{lk}'^* B'_{lk}\right)^{-1} B_{jk}'^* U_k.$$

Since $B_{kk} = I_{L_k}$ and $B'_{kk} = I_{L'_k}$, setting $i = j = k$ in (2) yields

$$(3) \quad U_k\left(\sum_{l=1}^{k} B_{lk}^* B_{lk}\right)^{-1} = \left(\sum_{l=1}^{k} B_{lk}'^* B'_{lk}\right)^{-1} U_k.$$

Setting $j = k$ in (2) and applying (3), we get

$$U_i B_{ik}\left(\sum_{l=1}^{k} B_{lk}^* B_{lk}\right)^{-1} = B'_{ik}\left(\sum_{l=1}^{k} B_{lk}'^* B'_{lk}\right)^{-1} U_k$$

$$= B'_{ik} U_k\left(\sum_{l=1}^{k} B_{lk}^* B_{lk}\right)^{-1},$$

and consequently

$$(4) \quad U_i B_{ik} = B'_{ik} U_k.$$

(b) Since

$$H = \bigoplus_{i=1}^{n} L_i = \bigoplus_{i=1}^{n} L'_i,$$

it is clear that if $U_i:L_i \to L'_i$ are unitary operators, then there exists a unique unitary operator U on H whose restriction to L_i is U_i. Suppose, further, that the U_i's satisfy (4). Taking adjoints yields

$$B^*_{ik}U^*_i = U^*_k B'^*_{ik};$$

multiplying this equation on the left and right by U_k and U_i, respectively, we get

(5) $U_k B^*_{ik} = B'^*_{ik} U_i.$

Hence,

$$U_k B^*_{ik} B_{ik} = B'^*_{ik} U_i B_{ik} = B'^*_{ik} B'_{ik} U_k,$$

for each i, whence,

$$U_k \left(\sum_{l=1}^{k} B^*_{lk} B_{lk} \right) = \left(\sum_{l=1}^{k} B'^*_{lk} B'_{lk} \right) U_k;$$

inversion now gives

$$\left(\sum_{l=1}^{k} B^*_{lk} B_{lk} \right)^{-1} U^*_k = U^*_k \left(\sum_{l=1}^{k} B'^*_{lk} B'_{lk} \right)^{-1};$$

pre and post multiplying this last equation by U_k yields equation (3). A successive application of equations (4), (3) and (5) to the left side of equation (2) shows that equation (2) is valid. Hence, we have shown that

$$U_i C_{k,ij} = C'_{k,ij} U_j,$$

where, of course, $(C'_{k,ij})$ the matrix of P'_k in the decomposition

$$H = \bigoplus_{i=1} L'_i.$$

It follows at once that $UP_k = P'_k U$, or, equivalently, that $U(M_k) = M'_k$ for each k.

Notation. If $M_1, \ldots, M_n$, $L_1, \ldots, L_n$ and the B_{jk}'s are as in Theorem 1, let B be the operator on H given, with respect to the decomposition

$$H = \bigoplus_{j=1}^{n} L_j,$$

by the upper-triangular operator matrix

$$(6)\quad B = \begin{bmatrix} 0 & B_{12} & B_{13} \dots & B_{1n} \\ 0 & 0 & B_{23} \dots & B_{2n} \\ 0 & 0 & 0 \quad \dots & B_{3n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 \quad \dots & 0 \end{bmatrix}$$

Theorem 2 has the following obvious reformulation: If $\{M_1, \dots, M_n\}$ and $\{M'_1, \dots, M'_n\}$ are two n-tuples of subspaces, both yielding algebraic direct sum decomposition of H, if B and B' are the operator matrices associated to the two n-tuples via (6), then the n-tuples $(M_1, \dots, M_n)$ and $(M'_1, \dots, M'_n)$ are unitarily equivalent if and only if the matrices B and B' are unitarily equivalent via a 'block-diagonal' unitary matrix.

Since it would be desirable, if possible, to identify the unitary equivalence problem for the n-tuple $(M_1, \dots, M_n)$ with the unitary equivalence problem for the associated B-operator, we shall now investigate the condition of block-diagonality of a unitary operator intertwining two B-operators.

LEMMA 4. *Let B be the operator matrix given by* (6). *Suppose $B_{k-1,k}$ is one-to-one, for $1 < k \leqq n$. Then,*

$$\ker B^k = L_1 \oplus \dots \oplus L_k, \quad \textit{for } 1 \leqq k \leqq n.$$

Proof. First consider ker B. Let $Bx = 0$, where x is given by the column vector $x = (x_1, \dots, x_n)'$ (the prime denoting transpose). Then, for $1 \leqq j \leqq n - 1$,

$$\sum_{k=j+1}^{n} B_{jk}x_k = 0.$$

For $j = n - 1$, this is $B_{n-1,n}x_n = 0$, which implies $x_n = 0$, by the assumed injectivity. If, inductively, it has been shown that $x_n = \dots = x_{j+2} = 0$, the above equation becomes

$$B_{j,j+1}x_{j+1} = 0,$$

which again forces $x_{j+1} = 0$. Thus, we conclude that $x_2 = \dots = x_n = 0$, or in other words, that ker $B = L_1$.

To discuss the case $k > 1$, the following bit of terminology will help; for any $n \times n$ matrix (A_{ij}) and $1 \leqq j \leqq n$, let us call $(A_{1j}, A_{2,j+1}, \dots, A_{n-j+1,n})$ the j-th diagonal of the matrix. Thus, for instance, the matrix

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

has (1, 5, 9), (2, 6) and (3) as its first, second and third diagonals.

Coming back to the proof, fix a k, with $1 \leqq k \leqq n$. It is not hard to show (by induction, for instance) that (a) the first k diagonals of B^k are identically zero; and (b) the $(k + 1)$-st diagonal of B^k is $(B_{12}B_{23} \dots B_{k,k+1}, B_{23} \dots B_{k+1,k+2}, \dots, B_{n-k+1,n-k+2} \dots B_{n-1,n})$. The hypothesis ensures now that every entry in this diagonal is an injective operator. Now, arguing exactly as in the case $k = 1$, it may be shown that

$$\ker B^k = L_1 \oplus \dots \oplus L_k.$$

The relationship between the M_k's and the B_{kj}'s reveals that injectivity of $B_{k-1,k}$ is equivalent to the condition

$$(M_1 + \dots + M_{k-2} + M_k) \cap (M_1 + \dots + M_{k-1})^{\perp} = \{0\}.$$

This prompts the following definition.

Definition 1. The ordered n-tuple $(M_1, \dots, M_n)$ of closed subspaces of H is said to be *generic* if, for $1 < k \leqq n$,

$$(M_1 + \dots + M_{k-2} + M_k) \cap (M_1 + \dots + M_{k-1})^{\perp} = (0).$$

Remark 2. (a) For $n = 2$, this gives only one condition:

$$M_1^{\perp} \cap M_2 = \{0\}.$$

This is a weaker condition than the one defined by Halmos (cf. [**4**]; he calls a pair (M_1, M_2) of subspaces to be in generic position if

$$M_1 \cap M_2 = M_1^{\perp} \cap M_2 = M_1 \cap M_2^{\perp} = M_1^{\perp} \cap M_2^{\perp} = \{0\}.$$

For one thing, his notion is a symmetric one; i.e., the order in the pair (M_1, M_2) is irrelevant. It is not hard to see that, for finite dimensional H, a pair of subspaces (M_1, M_2) is in generic position in the sense of Halmos if and only if (i) H is the algebraic direct sum of M_1 and M_2, and (ii) both the ordered pairs (M_1, M_2) and (M_2, M_1) are generic in the sense of Definition 1 above.

(b) If

$$H = \sum_{i=1}^{n} M_i$$

is an algebraic direct sum, and if the operators B_{jk} are constructed as in Theorem 1, then, genericity of $(M_1, \dots, M_n)$ is equivalent to injectivity of

https://doi.org/10.4153/CJM-1988-002-0 Published online by Cambridge University Press

each $B_{k-1,k}$. In particular, if $\dim H < \infty$, then $\dim M_i = \dim M_j$ for all j, and $\dim H = n \dim M_1$.

(c) The term 'generic' is apt, in the following sense: if

$$H = \sum_{i=1}^{n} M_i$$

is an algebraic direct sum, if $\dim H < \infty$ and $\dim M_i = \dim M_j$ for all i, j, then, for any $\epsilon > 0$, there exists an algebraic direct sum decomposition

$$H = \sum_{i=1}^{n} M'_i$$

such that $(M'_1, \ldots, M'_n)$ is generic and

$$\|P_i - P'_i\| < \epsilon \quad \text{for } 1 \leqq i \leqq n,$$

where P_i and P'_i are the orthogonal projections onto M_i and M'_i respectively. (Reason: if

$$H = \bigoplus_{i=1}^{n} L_i$$

is the orthogonal direct sum decomposition associated with

$$H = \sum_{i=1}^{n} M_i,$$

and if $\{B_{jk}:1 \leqq j < k \leqq n\}$ are the operators given by Theorem 1, let $(M'_1, \ldots, M'_n)$ be the n-tuple determined by the orthogonal decomposition

$$H = \bigoplus_{i=1}^{n} L_i$$

and the operators $\{B'_{jk}:1 \leqq j < k \leqq n\}$, where $B'_{jk} = B_{jk}$ if $j < k - 1$, and $B'_{k-1,k}$ is an invertible operator from L_k to L_{k-1} such that

$$\|B_{k-1,k} - B'_{k-1,k}\| < \delta \quad \text{for all } k,$$

where δ is chosen small enough to ensure

$$\|P_k - P'_k\| < \epsilon;$$

this is possible by the representations of P_k and P'_k given by Lemma 3).

(d) The observation in (c) above can be strengthened to the following more symmetric assertion (the proof being identical): with the notation of (c), one can choose the M'_i such that

$$\|P_i - P'_i\| < \epsilon \quad \text{for all } i,$$

and such that

$$\{M'_{\sigma(1)}, \dots, M'_{\sigma(n)}\}$$

is 'generic', for each permutation σ. Thus, the remark (c) is not meant as a justification for the asymmetry of Definition 1; that justification and, in fact, the raison d'être of Definition 1 lies in the next proposition, where the reader may observe that genericity plays a crucial role, and is in fact, quite close to being a necessary condition (though not quite) for the validity of the assertion.

THEOREM 3. *Let*

$$H = \sum_{i=1}^{n} M_i = \sum_{i=1}^{n} M'_i$$

be two algebraic direct sum decompositions of H. Suppose both the n-tuples $(M_1, \dots, M_n)$ and $(M'_1, \dots, M_n)$ are generic. Let B and B' be the operators associated to these n-tuples via equation (6). *For a unitary operator U on H, the following conditions are equivalent*:

(i) $U(M_i) = M'_i$ *for* $1 \leqq i \leqq n$;

(ii) $UBU^* = B'$.

Proof. The implication (i) $\Rightarrow$ (ii) is a direct consequence of Theorem 2. For the converse implication, it suffices (again, by Theorem 2) to prove that any U as in (ii) must be in 'block-diagonal' form, i.e.; we must show that if $UBU^* = B'$, then U must necessarily map L_k onto L'_k for $1 \leqq k \leqq n$. However, if $UBU^* = B'$, then it is clear that

$$U(\ker B^k) = \ker B'^k.$$

By Lemma 4, this says that

$$U(L_1 \oplus \dots \oplus L_k) = L' \oplus \dots \oplus L'_k$$

for each k. Since the L_i's (respectively, the L'_i's) are mutually orthogonal subspaces, this ensures that $U(L_k) = L'_k$ for all k, as desired.

3. Generators of $\mathcal{B}(H)$. For any subset S of $\mathcal{B}(H)$, let us write $W^*(S)$ for the von-Neumann algebra generated by S. In [**1**], Davis shows that (a) if P_1 and P_2 are orthogonal projections on a Hilbert space H with $\dim H > 2$, then

$$W^*(\{P_1, P_2\}) \subsetneq \mathcal{B}(H);$$

while (b) if H is a separable infinite-dimensional Hilbert space, there exist three orthogonal projections P_1, P_2 and P_3 on H such that

$$W^*(\{P_1, P_2, P_3\}) = \mathcal{B}(H).$$

It will be shown below, using the results of the preceding sections, that if $n \geqq 3$, and if H is a separable Hilbert space which is either infinite dimensional or finite-dimensional with dimension a multiple of n, there exist n orthogonal projections $P_1, \ldots, P_n$ on H such that

(i) $W^*(\{P_1, \ldots, P_n\}) = \mathscr{B}(H)$ and

(ii) $W^*(S) \neq \mathscr{B}(H)$ whenever $S \subsetneq \{P_1, \ldots, P_n\}$.

THEOREM 4. *Let L be a separable Hilbert space and let H be the Hilbert space direct sum of n copies of L. Let*

$$\{B_{jk}: 1 \leqq j < k \leqq n\} \subseteq \mathscr{B}(L)$$

satisfy (a) B_{in} *has dense range, for* $1 \leqq i < n$; *and*

(b) $W^*(\{B_{in}^* B_{in}: 1 \leqq i < n\}) = \mathscr{B}(L)$.

(If $n \geqq 3$, *these conditions can be met by an appropriate choice of the* B_{jk}*'s). Then, if* P_k *denotes the orthogonal projection onto the subspaces* M_k *of* H *defined by*

$$M_k = \{(B_{1k}x, \ldots, B_{k-1,k}x, x, 0, \ldots, 0): x \in L\},$$

the following assertions hold:
(i) $W^*(\{P_1, \ldots, P_n\}) = \mathscr{B}(H)$;
(ii) $W^*(S) \subsetneq \mathscr{B}(H)$, *whenever* $S \subsetneq \{P_1, \ldots, P_n\}$.

Proof. First, let us prove the parenthetical statement which ensures that the above theorem is not a vacuous statement. To see this, note first that since L is separable, there exists $C \in \mathscr{B}(L)$ such that $W^*(\{C\}) = \mathscr{B}(L)$. (For example, if $\dim L = \aleph_0$, so that L may be taken as l^2, we may take C to be the unilateral shift; if $\dim L = m$, and if $\{e_1, \ldots, e_m\}$ is an orthonormal basis for L, let C be the operator defined by $Ce_m = 0$, $Ce_i = e_{i+1}$, for $1 \leqq i < m$.) Let $C = A_1 + iA_2$ be the cartesian decomposition of C. Define $B_{jk} = I$, if $1 \leqq j < k < n$, or if $2 < j < n$ (this is where $n > 2$ is required), and define

$$B_{jn} = [A_j + 2\|A_j\|I]^{1/2} \quad \text{for } j = 1, 2.$$

This choice of B_{jk}'s satisfies conditions (a) and (b).

For the proof of the theorem, if M_k is defined via the B_{jk}'s as above, then

$$H = \sum_{i=1}^{n} M_i$$

is an algebraic direct sum (cf. Remark 1 (b)). It is clear that if

$$S \subsetneq \{P_1, \ldots, P_n\},$$

then $\sum \{M_i : P_i \in S\}$ is a non-trivial invariant subspace for each P_i in S, so that, by the double commutant theorem, $W^*(S)$ must be properly contained in $\mathscr{B}(E)$; thus (ii) is established.

In order to establish (i), since each P_i is self-adjoint, it suffices, in view of the double commutant theorem and the fact that any C^*-algebra (in this case, the commutant of $W^*(\{P_1, \dots, P_n\})$) is linearly spanned by its unitary elements, to show that if U is a unitary operator on H such that $UP_k = P_kU$ for all k, then $U = \omega I$ for some complex number ω of unit modulus. So, suppose U is a unitary operator on H such that $UP_k = P_kU$ for all k. Clearly then, $U(M_k) = M_k$ for all k. It follows from Theorem 2 (choosing $M_i' = M_i$) that with respect to the decomposition $H = L \oplus \dots \oplus L$, U has a block-diagonal matrix $U = \operatorname{diag}(U_i)$. Theorem 2 then asserts that

$$U_iB_{ik} = B_{ik}U_k \quad \text{for } 1 \leqq i < k \leqq n.$$

Exactly as in the proof of Theorem 2, it may now be deduced that

$$U_nB_{in}^*B_{in} = B_{in}^*B_{in}U_n \quad \text{for } 1 \leqq i < n;$$

i.e.,

$$U_n \in \{B_{in}^*B_{in} : 1 \leqq i < n\}'.$$

It follows from hypothesis (b) and the double commutant theorem that $U_n = \omega I_L$ for some complex number ω of unit modulus. Then, the equation

$$U_iB_{in} = B_{in}U_n = \omega B_{in}$$

and the hypothesis (a) guarantees that $U_i = \omega I_L$ for each i; in other words $U = \omega I_H$, as desired.

4. Canonical (partial) correlation coefficients. Let $X_1, \dots, X_p$ and $Y_1, \dots, Y_q$ be two sets of random variables on a probability space $(\Omega, \mathscr{B}, P)$, each with finite variance and mean zero. Hotelling proposed (in **[5]**) the 'canonical correlation coefficients' as a measure of the strength of linear association between the two sets of random variables, as follows:

Let M (respectively, N) be the space of linear combinations of the X_i's (respectively, the Y_i's). (Then, of course, by the assumed existence of finite variances, the spaces M and N are linear subspaces of $L^2(\Omega, \mathscr{B}, P)$. In the sequel, the inner product and norm used will be the ones on $L^2(P)$; thus, $\langle X, Y\rangle = E(X\bar{Y})$. (In the real case, of course, there is no need for complex conjugation).) Define

$$\rho_1 = \sup\{ |\langle X, Y\rangle| : X \in M,\ Y \in N,\ \|X\| = 1 = \|Y\| \}.$$

Pick X_1' in M and Y_1' in N such that

$$||X_1'|| = ||Y_1'|| = 1 \quad \text{and} \quad \langle X_1', Y_1'\rangle = \rho_1.$$

Next, let

$$M_1 = \{X \in M: \langle X, X_1'\rangle = 0\} \quad \text{and}$$

$$N_1 = \{Y \in N: \langle Y, Y_1'\rangle = 0\},$$

and define

$$\rho_2 = \sup\{ |\langle X, Y\rangle| : X \in M_1, Y \in N_1, ||X|| = 1 = ||Y|| \}.$$

Pick X_2' in M_1 and Y_2' in N_1 such that

$$||X_2'|| = ||Y_2'|| = 1 \quad \text{and} \quad \langle X_2', Y_2'\rangle = \rho_2.$$

Then, let

$$M_2 = \{X \in M: \langle X, X_i'\rangle = 0 \text{ for } 1 \leqq i \leqq 2\} \quad \text{and}$$

$$N_2 = \{Y \in N: \langle Y, Y_i'\rangle = 0 \text{ for } 1 \leqq i \leqq 2\},$$

and pick X_3' in M_2 and Y_3' in N_2 such that

$$||X_3'|| = 1 = ||Y_3'|| \quad \text{and} \quad \langle X_3', Y_3'\rangle = \rho_3,$$

where

$$\rho_3 = \sup\{ |\langle X, Y\rangle| : X \in M_3, Y \in N_3, ||X|| = 1 = ||Y|| \}.$$

Continuing this process to its logical conclusion results in sequences $\{\rho_1, \dots, \rho_k\}$, $\{X_1', \dots, X_k'\}$ and $\{Y_1', \dots, Y_k'\}$, where k is the minimum of dim M and dim N. The non-zero ρ_i's are called the canonical correlation coefficients (they do not depend on the choice of the X_i''s and Y_i''s) and the X_i''s and Y_i''s are called the canonical variables.

This notation was extended by Roy (in [**9**]) to three sets of random variables as follows: Let $X_1, \dots, X_p$; $Y_1, \dots, Y_q$; $Z_1, \dots, Z_r$ be three sets of random variables of finite variance and mean zero. Let M_1, M_2, M_3 denote the linear spaces spanned by these sets, respectively. Roy defined the canonical partial correlation coefficients between $\{Y_1, \dots, Y_q\}$ and $\{Z_1, \dots, Z_r\}$ as the canonical correlation coefficients between $\{\tilde{Y}_1, \dots, \tilde{Y}_q\}$ and $\{\tilde{Z}_1, \dots, \tilde{Z}_r\}$, where

$$Y_i = Y_i - P_1(Y_i) \quad \text{and} \quad \tilde{Z}_i = Z_i - P_1(Z_i),$$

the symbol P_1 denoting the orthogonal projection (in L^2) onto M_1.

In this section, we shall apply Theorem 1 to the problem of determining these correlation coefficients.

(a) *Canonical correlation coefficients*. Let $X_1, \dots, X_p$ and $Y_1, \dots, Y_q$ be two collections of random variables of finite variance and mean zero. Let M and N denote the linear spaces spanned by them, respectively, and let $H = M + N$ (equipped with the inner product coming from L^2).

Case (i). $M \cap N = (0)$. In this case, $H = M + N$ is an algebraic direct sum decomposition, and so, by Theorem 1, there exists a linear operator $B:M^{\perp} \to M$ (in case $n = 2$), we have $L_1 = M$, $L_2 = M^{\perp}$) such that

$$N = \{ (By, y):y \in M^{\perp}\}$$

with respect to the decomposition $H = M \oplus M^{\perp}$. Then, by definition,

$$\begin{aligned}
\rho_1 &= \sup\{ | \langle (x, 0), (By, y) \rangle :x \in M, y \in M^{\perp}, \|x\|^2 = 1 \\
&= \|By\|^2 + \|y\|^2 \\
&= \sup\{ | \langle (x, By\rangle |:x \in M, y \in M^{\perp}, \|x\|^2 = 1 \\
&= \|By\|^2 + \|y\|^2\} \\
&= \sup\{ \|By\|:y \in M^{\perp}, \|y\|^2 + \|By\|^2 = 1\} \\
&= \sup\{ \|By\|:y \in M^{\perp}, \langle (I + B^*B)y, y\rangle = 1\} \\
&= \sup\{ \|By\|:y \in M^{\perp}, \| (I + B^*B)^{1/2}y\| = 1\} \\
&= \sup\{ \|B(I + B^*B)^{-1/2}z\|:z \in M^{\perp}, \|z\| = 1\}.
\end{aligned}$$

It follows, by a successive application of arguments similar to the ones used in obtaining the above string of equalities, that if $\{Y'_1, \dots, Y'_l\}$ is an orthonormal basis for $M^{\perp}$ such that

$$B^*By'_i = \alpha_i^2 y'_i,$$

with $\alpha_1 \geqq \dots \geqq \alpha_l \geqq 0$, then the canonical variables are given by

$$\left\{\frac{1}{\alpha_i}(By'_i, 0): i = 1, 2, \dots, k\right\} \quad \text{and}$$

$$\left\{\frac{1}{(1 + \alpha_i^2)^{1/2}}(By'_i, y'_i): i = 1, 2, \dots, k\right\}$$

while the canonical correlation coefficients are given by

$$\rho_i = \alpha_i(1 + \alpha_i^2)^{-1/2}, \quad \text{for } i = 1, \dots, k,$$

where k is the rank of B.

Case (ii). $M \cap N \neq (0)$. Let

$$M' = M \cap (M \cap N)^{\perp}, N' = N \cap (M \cap N)^{\perp} \quad \text{and}$$

$$H' = M' + N'.$$

Then $H' = M' + N'$ is an algebraic direct sum; let

$$B':M'^{\perp} \to M'$$

(here, the orthogonal complement is taken relative to H') such that, in the decomposition $H' = M' \oplus M'^{\perp}$ the subspace N' is described by

$$N' = \{ (B'y, y) : y \in M^{\perp} \}.$$

If the singular values of B' are $\alpha_1, \ldots, \alpha_l$ (written in decreasing order), and if k' is the rank of B', it is not hard to see, using case (i), that the canonical correlation coefficients are given by

$$\{1, 1, \ldots, 1, \alpha_i(1 + \alpha_i^2)^{-1}, \ldots, \alpha_{k'}(1 + \alpha_{k'}^2)^{-1}\},$$

where the length of the initial string of 1's is equal to dim $(M \cap N)$.

(b) *Canonical partial correlation coefficients*. Let $X_1, \ldots, X_p$; $Y_1, \ldots, Y_q$; $Z_1, \ldots, Z_r$ be three sets of random variables of finite variance and mean zero. Let M_1, M_2, M_3 be the linear spaces spanned by the three sets, respectively. We shall compute the canonical partial correlation coefficients between the Y and Z sets. Let $H = M_1 + M_2 + M_3$.

Case (i): $H = \sum_{i=1}^{3} M_i$ is an algebraic direct sum. Let the spaces L_1, L_2, L_3 and the operators $B_{jk}(1 \leqq j < k \leqq 3)$ be constructed as in Theorem 1. Since the projection onto $M_1^{\perp}$ sends M_2 and M_3 to the subspaces $\widetilde{M}_2$ and $\widetilde{M}_3$ of $L_2 \oplus L_3$ given by

$$\widetilde{M}_2 = \{ (y, 0) : y \in L_2\} \quad \text{and} \quad \widetilde{M}_3 = \{ (B_{23}z, z) : z \in L_3\},$$

it can be shown, exactly as in Case (i) of (a), that the canonical partial correlation coefficients between the Y_i's and the Z_i's are given by

$$\rho_i = \alpha_i(1 + \alpha_i^2)^{-1}, \quad 1 \leqq i \leqq k,$$

where k is the rank of B_{23} and $(\alpha_1, \ldots, \alpha_l)$ is an enumeration, in decreasing order, of the singular values of B_{23}.

Case (ii). $H = \sum_{i=1}^{3} M_i$ is not an algebraic direct sum. It is easy to see that

$$H = \sum_{i=1}^{3} M_i$$

is an algebraic direct sum if and only if

$$M_1 \cap M_2 = \{0\} = (M_1 + M_2) \cap M_3.$$

It is, hence, natural in this case to define the subspaces

$$M_2' = M_2 \cap (M_2 \cap M_1)^{\perp} \quad \text{and}$$

$$M_3' = M_3 \cap (M_3 \cap (M_1 + M_2))^{\perp}.$$

It is clear that $M_1 + M_2 = M_1 + M_2'$ and that $H = M_1 + M_2' + M_3'$ is an algebraic direct sum decomposition. Apply Theorem 1 to the subspaces M_1, M_2', M_3' to get an orthogonal decomposition

$$H = L_1 \oplus L_2 \oplus L_3$$

https://doi.org/10.4153/CJM-1988-002-0 Published online by Cambridge University Press

and the operators

$$B_{jk}:L_k \to L_j \quad \text{for } 1 \leqq j < k \leqq 3.$$

It is not too hard then to show that if $\alpha_1 \geqq \ldots \geqq \alpha_l \geqq 0$ are the singular values of B_{23}, then the canonical partial correlation coefficients of $\{Y_1, \ldots, Y_q\}$ and $\{Z_1, \ldots, Z_r\}$ are given by

$$1, 1, \ldots, 1, \alpha_1(1 + \alpha_1^2)^{-1}, \ldots, \alpha_m(1 + \alpha_m^2)^{-1},$$

where the length of the initial string of 1's is equal to

$$\dim((M_1 + M_2) \cap (M_1 + M_3) \cap M_1^{\perp}),$$

and m is the rank of B_{23}.

It may be advisable to point out that replacing B_{23} by B_{12} (for instance) in the above discussion would not lead to the canonical partial correlation coefficients between the X_i's and the Y_i's. To apply the above procedure, the span of the set of random variables, whose linear effect is to be ignored, must be taken as M_1, while the second and third subspaces must be taken as the spans of the sets of random variables whose canonical partial correlation coefficients are to be computed.

References

1. C. Davis, *Generators of the ring of bounded operators*, Proc. Amer. Math. Soc. *6* (1955), 970-972.
2. ——— *Separation of two linear subspaces*, Acta Sci. Math. Szeged. *19* (1958), 172-187.
3. C. Davis and W. M. Kahan, *Rotation of eigenveetors*, SIAM J. Numer. Anal. *7* (1970), 1-16.
4. P. R. Halmos, *Two subspaces*, Trans. Amer. Math. Soc. *144* (1969), 381-389.
5. H. Hotelling, *Relations between two sets of variables*, Biometrika *28* (1936), 321-377.
6. C. Jordan, *Essai sur la geometrie a n dimensions*, Bull. Soc. Math. France *3* (1875), 103-174.
7. T. Kato, *Perturbation theory for linear operators* (Springer, Berlin, 1966).
8. F. J. Murray and J. von Neumann, *On rings of operators*, Ann. of Math. *37* (1936), 116-229.
9. S. N. Roy, *Some aspects of multivariate analysis* (Wiley, New York, 1958).

Indian Statistical Institute,
New Delhi, India