COMPRESSIONS TO FINITE-DIMENSIONAL SUBSPACES

CHANDLER DAVIS[1]

Let $\mathcal{H}$ be hilbert space (any dimensionality, real or complex scalars). Let P be a hermitian projection. Let A be any hermitian operator. The *compression* of A to $P\mathcal{H}$ [**3**] is PAP, considered as an operator on $P\mathcal{H}$. Compressions of completely continuous positive operators are of interest in connection with estimating eigenvalues: the Fischer-Courant minimax theorem [**5**, p. 235] says the kth highest eigenvalue of PAP is not greater than that of A. Compressions enter in the study of more general mappings of operators, often via Naĭmark's theorem [**6**].

Especially in the first connection, the case where $P\mathcal{H}$ is finite-dimensional is interesting. But in some problems a finite-dimensional subspace may be known, not via the operator P, but via an arbitrary set of vectors which span it; if they are not orthonormal, one would rather not have to find P. This suggests that the following elementary formulas may be worth pointing out. I suppose that at least Formula 1 must be known already, but not, apparently, very widely.

NOTATION. *$x_1, \cdots, x_n$ form a linear basis of $P\mathcal{H}$. G denotes the determinant of their Gramian ($n \times n$ matrix with i, j entry (x_i, x_j)). If the kth row of the Gramian is replaced by $(z, x_1), \cdots, (z, x_n)$, all other rows being left unchanged, the determinant of the resulting matrix will be denoted $G(x_k; z)$.*

Evident properties: $G(x_k; z) = 0$ if $z = (1-P)z$ or $z = x_i$ $(i \neq k)$, while $G(x_k; x_k) = G$; also $G(x_k; z)$ is linear in z. These may be summed up by saying that $G(x_k; z) = (z, x_k^*)G$, where $\{x_1^*, \cdots, x_n^*\}$ is the basis of $P\mathcal{H}$ biorthonormal with $\{x_1, \cdots, x_n\}$.

FORMULA 1. $Pz = G^{-1} \sum_k G(x_k; z)x_k$. (This notation here and below means summation over all available values of the index.)

PROOF. Uniquely $z = \sum_i a_i x_i + (1-P)z$. Substitute this on both sides, and use the evident properties of $G(x_k; z)$.

FORMULA 2. $\mathrm{tr}(PAP) = G^{-1} \sum_k G(x_k; Ax_k)$.

PROOF. Let $\xi_1, \cdots, \xi_n$ be orthonormal eigenvectors of PAP, and $\lambda_1, \cdots, \lambda_n$ their respective eigenvalues; then $x_i = \sum_\rho T_{i\rho}\xi_\rho$, where T is some nonsingular matrix. Recall that

$$(x_i, x_j) = \sum_{\rho\sigma} T_{i\rho}\overline{T}_{j\sigma}(\xi_\rho, \xi_\sigma) = \sum_\rho T_{i\rho}\overline{T}_{j\rho} = (TT^*)_{ij}$$

Received by the editors December 6, 1957.

[1] National Science Foundation Fellow.

gives $G = |\det T|^2$. (* denotes conjugate transpose.) It remains to prove that, analogously, $\sum_k G(x_k;\ Ax_k) = |\det T|^2 \operatorname{tr}(PAP)$. Now $(Ax_i,\ x_j) = (PAPx_i,\ x_j) = \sum_\rho T_{i\rho}\lambda_\rho \overline{T}_{j\rho}$. So if we let R^k denote the matrix with entries $R^k_{i\rho} = T_{i\rho}$ $(i \neq k)$ and $R^k_{k\rho} = T_{k\rho}\lambda_\rho$, we have $G(x_k; Ax_k) = \det(R^kT^*)$, $\sum_k G(x_k;\ Ax_k) = \det T^* \sum_k \det R^k$. In the expansion

$$\sum_k \det R^k = \sum_k \sum_{\rho_1, \cdots, \rho_n} \epsilon^{1 \cdots n}_{\rho_1, \cdots, \rho_n} T_{1\rho_1} \cdots T_{n\rho_n} \lambda_{\rho_k}$$

the summation over k may be carried out first: $\sum_k \lambda_{\rho_k} = \operatorname{tr}(PAP)$ for any $(\rho_1, \cdots, \rho_n)$ giving a nonzero contribution. This gives the result.

The proof would have been simpler if I had exploited the evident properties of $G(x_k; z)$. I gave this version because Formula 3 is proved altogether analogously, without introducing any new notions.

Instead of the trace c_1, consider now c_ν, where for any B

$$\det(\lambda + B) = \sum_\nu c_\nu(B)\lambda^{n-\nu};$$

that is, c_ν is the νth elementary symmetric polynomial of the eigenvalues. Extend the notation: $G(x_{k_1};\ z_1)(x_{k_2};\ z_2)$ is the determinant of the matrix which differs from the Gramian in having k_1th row $(z_1,\ x_1), \cdots, (z_1,\ x_n)$ *and* in having k_2th row $(z_2,\ x_1), \cdots, (z_2,\ x_n)$; and so forth.

FORMULA 3.[2] $c_\nu(PAP) = G^{-1} \sum G(x_{k_1};\ Ax_{k_1}) \cdots (x_{k_\nu};\ Ax_{k_\nu})$. (In this equation summation is over all distinct ν-tuples $\{k_1, \cdots, k_\nu\}$ from among $\{1, \cdots, n\}$.)

PROOF. See under Formula 2.

An interesting case is where A is another projection Q. A complete set of unitary-invariants for the *pair* of subspaces $P\mathcal{H}$ and $Q\mathcal{H}$ is the spectrum of PQP and its multiplicity function (together with the dimensionalities of $Q\mathcal{H} \cap (1-P)\mathcal{H}$ and $(1-Q)\mathcal{H} \cap (1-P)\mathcal{H}$) [1;2].[3] For a simple numerical measure of the closeness of $P\mathcal{H}$ to being contained in $Q\mathcal{H}$, $\operatorname{tr}(PQP)$ recommends itself (or, if you like, $n^{-1}\operatorname{tr}(PQP)$). If $Q\mathcal{H}$ is finite-dimensional, one may ask for a modification of Formula 2 which treats P and Q symmetrically.

$y_1, \cdots, y_m$ form a linear basis of $Q\mathcal{H}$. H denotes the determinant of their Gramian; $H(y_l;\ z)$, etc. are defined in analogy to previous notations.

FORMULA 4. $\operatorname{tr}(PQP) = \operatorname{tr}(QPQ) = (GH)^{-1} \sum_{kl} G(x_k;\ y_l) H(y_l;\ x_k)$.

[2] The case $\nu = n$ shows the equivalence of Theorem 1 of [4] to Weyl's theorem which it generalizes.

[3] One might prefer replacing PQP by $PQP + (1-P)(1-Q)(1-P) = 1 - P - Q + PQ + QP$, making apparent the symmetrical roles of P and Q [1].

PROOF. I have proofs of Formulas 4 and 5 along the unsophisticated lines followed above for Formulas 2 and 3, but they are clumsy. Instead, rewrite the right side of Formula 4 in terms of the biorthonormal bases $\{x_1, \cdots, x_n\}$, $\{x_1^*, \cdots, x_n^*\}$ of $P\mathcal{H}$, $\{y_1, \cdots, y_m\}$, $\{y_1^*, \cdots, y_m^*\}$ of $Q\mathcal{H}$. It equals

$$\sum_{kl} (y_l, \overset{*}{x}_k)(x_k, \overset{*}{y}_l) = \sum_{kl} ((x_k, \overset{*}{y}_l)y_l, \overset{*}{x}_k) = \sum_k (Qx_k, \overset{*}{x}_k) = \mathrm{tr}(PQP),$$

by Formulas 1 and 2.

FORMULA 5.

$$c_\nu(PQP) = c_\nu(QPQ)$$
$$= (GH)^{-1} \sum G(x_{k_1}; y_{l_1}) \cdots (x_{k_\nu}; y_{l_\nu}) H(y_{l_1}; x_{k_1}) \cdots (y_{l_\nu}; x_{k_\nu}).$$

(In this equation summation is over all distinct pairs of ν-tuples, $\{k_1, \cdots, k_\nu\}$ from among $\{1, \cdots, n\}$ and $\{l_1, \cdots, l_\nu\}$ from among $\{1, \cdots, m\}$.)

PROOF. The equation

$$G(x_{k_1}; z_1) \cdots (x_{k_\nu}; z_\nu) = G\nu!(z_1 \otimes \cdots \otimes z_\nu, \overset{*}{G}_{k_1\cdots k_\nu})$$

defines an element $G^*_{k_1\cdots k_\nu}$ of $\mathcal{H}^\nu$, the tensor product of ν copies of $\mathcal{H}$. Extend $\{x_1, \cdots, x_n\}$ to a basis of $\mathcal{H}$ by adjoining an orthonormal basis $\{x_{n+1}, x_{n+2}, \cdots\}$ of $(1-P)\mathcal{H}$. The elements $x_{\sigma_1} \otimes \cdots \otimes x_{\sigma_\nu}$ form a linear basis of $\mathcal{H}^\nu$. By considering its scalar products with these basis vectors, $G^*_{k_1\cdots k_\nu}$ is identified as

$$\frac{1}{\nu!} \sum_{l_1\cdots l_\nu} \epsilon^{k_1\cdots k_\nu}_{l_1\cdots l_\nu} \overset{*}{x}_{l_1} \otimes \cdots \otimes \overset{*}{x}_{l_\nu} = \overset{*}{x}_{[k_1} \otimes \cdots \otimes \overset{*}{x}_{k_\nu]}.$$

(Again ϵ is defined by $\epsilon = \pm 1$ if $(l_1, \cdots, l_\nu)$ is respectively an even or an odd permutation of $(k_1, \cdots, k_\nu)$, $\epsilon = 0$ otherwise. The bracket on the subscripts, denoting antisymmetrization, is defined by the equation.)

The easily-proved analog of Formula 1 is

$$Pz_{[1} \otimes \cdots \otimes Pz_{\nu]}$$
$$= \sum_{k_1\cdots k_\nu} (z_1 \otimes \cdots \otimes z_\nu, \overset{*}{x}_{[k_1} \otimes \cdots \otimes \overset{*}{x}_{k_\nu]}) x_{k_1} \otimes \cdots \otimes x_{k_\nu}.$$

Formula 3 in the new notation reads

$$c_\nu(PAP) = \sum_{k_1\cdots k_\nu} (Ax_{k_1} \otimes \cdots \otimes Ax_{k_\nu}, \overset{*}{x}_{[k_1} \otimes \cdots \otimes \overset{*}{x}_{k_\nu]});$$

this is not disturbed if the subscripts of the Ax_{k_i} are also bracketed.

The right side of Formula 5 becomes

$$\sum_{k_1\cdots k_\nu;l_1\cdots l_\nu} (y_{l_1}\otimes\cdots\otimes y_{l_\nu}, \overset{*}{x}_{[k_1}\otimes\cdots\otimes \overset{*}{x}_{k_\nu]})(x_{k_1}\otimes\cdots\otimes x_{k_\nu}, \overset{*}{y}_{[l_1}\otimes\cdots\otimes \overset{*}{y}_{l_\nu]}).$$

By the analog of Formula 1 this is equal to

$$\sum_{k_1\cdots k_\nu} (Qx_{[k_1}\otimes\cdots\otimes Qx_{k_\nu]}, \overset{*}{x}_{[k_1}\otimes\cdots\otimes \overset{*}{x}_{k_\nu]}),$$

and by Formula 3 this is $c_\nu(PQP)$, as claimed.

The analogy to the special case, Formula 4, could be strengthened by mentioning that $Pz_{[1}\otimes\cdots\otimes Pz_{\nu]}=P_\nu(z_1\otimes\cdots\otimes z_\nu)$, where P_ν is the hermitian projection on the subspace of $\mathcal{H}^\nu$ linearly spanned by antisymmetrized products of elements of $P\mathcal{H}$. The $x_{[k_1}\otimes\cdots\otimes x_{k_\nu]}$ and the $x^*_{[k_1}\otimes\cdots\otimes x^*_{k_\nu]}$ are almost biorthonormal bases of $P_\nu\mathcal{H}^\nu$:

$$(x_{[l_1}\otimes\cdots\otimes x_{l_\nu]}, \overset{*}{x}_{[k_1}\otimes\cdots\otimes \overset{*}{x}_{k_\nu]}) = \frac{1}{\nu!}\,\epsilon^{l_1\cdots l_\nu}_{k_1\cdots k_\nu}.$$

References

1. C. Davis, *Separation of two linear subspaces*, Acta Sci. Math. Szeged., to appear.

2. J. Dixmier, *Position relative de deux variétés linéaires fermées dans nu sepace de Hilbert*, Revue Scientifique vol. 86 (1948) pp. 387–399.

3. P. Halmos, *Normal dilations and extensions of operators*, Summa Brasiliensis Mathematica vol. 2 (1950) fasc. 9.

4. A. Horn, *On the singular values of a product of completely continuous operators*, Proc. Nat. Acad. Sci. vol. 36 (1950) pp. 374–375.

5. F. Riesz and B. Sz.-Nagy, *Leçons d'analyse fonctionelle*, Budapest, 1954.

6. W. F. Stinespring, *Positive functions on C*-algebras*, Proc. Amer. Math. Soc. vol. 6 (1955) pp. 211–216.

INSTITUTE FOR ADVANCED STUDY