Let X be a 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 PX is \( PAP \), considered as an operator on PX. 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 Naimark’s theorem [6].

Especially in the first connection, the case where PX 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.** \( x_1, \ldots, x_n \) form a linear basis of PX. \( G \) denotes the determinant of their Gramian (an \( n \times n \) matrix with \( i, j \) entry \((x_i, x_j)\)). If the \( k \)th row of the Gramian is replaced by \((z, x_1), \ldots, (z, x_n)\), 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 \( \{x_1^*, \ldots, x_n^*\} \) is the basis of PX biorthonormal with \( \{x_1, \ldots, x_n\} \).

**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.** \( \text{tr}(PAP) = G^{-1}\sum_k G(x_k; A x_k) \).

**Proof.** Let \( \xi_1, \ldots, \xi_n \) be orthonormal eigenvectors of \( PAP \), and \( \lambda_1, \ldots, \lambda_n \) their respective eigenvalues; then \( x_i = \sum_{\rho} T_{i\rho} \xi_{\rho} \), where \( T \) is some nonsingular matrix. Recall that
\[
(x_i, x_j) = \sum_{\rho \sigma} T_{i\rho} \overline{T_{j\sigma}} (\xi_{\rho}, \xi_{\sigma}) = \sum_{\rho} T_{i\rho} \overline{T_{j\rho}} = (TT^*)_{ij}
\]
gives $G = |\det T|^2$. ($^*$ denotes conjugate transpose.) It remains to prove that, analogously, $\sum_k G(x_k; Ax_k) = |\det T|^2 \text{tr}(PAP)$. Now $(Ax_i, x_j) = (PAPx_i, x_j) = \sum_p T_{ip} \lambda_p T_{jp}$. So if we let $R^k$ denote the matrix with entries $R^k_{ip} = T_{ip} (i \neq k)$ and $R^k_{kp} = T_{kp} \lambda_p$, we have $G(x_k; Ax_k) = \det(R^k T^*)$, $\sum_k G(x_k; Ax_k) = \det T^* \sum_k \det R^k$. In the expansion

$$\sum_k \det R^k = \sum_k \sum_{p_1, \ldots, p_n} T_{1p_1} \cdots T_{np_n} \lambda_{p_k}$$

the summation over $k$ may be carried out first: $\sum_k \lambda_{p_k} = \text{tr}(PAP)$ for any $(p_1, \ldots, p_n)$ 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 $c_v$, where for any $B$

$$\det (\lambda + B) = \sum_v c_v(B)\lambda^{n-v};$$

that is, $c_v$ is the $v$th 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 $k_1$th row $(z_1, x_1), \ldots, (z_1, x_n)$ and in having $k_2$th row $(z_2, x_1), \ldots, (z_2, x_n)$; and so forth.

**Formula 3.** $c_v(PAP) = G^{-1} \sum G(x_{k_1}; Ax_{k_1}) \cdots (x_{k_v}; Ax_{k_v})$. (In this equation summation is over all distinct $v$-tuples $\{k_1, \ldots, k_v\}$ from among $\{1, \ldots, n\}$.)

**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 $P\mathcal{C}$ and $Q\mathcal{C}$ is the spectrum of $PQP$ and its multiplicity function (together with the dimensionalities of $Q\mathcal{C} \cap (1-P)\mathcal{C}$ and $(1-Q)\mathcal{C} \cap (1-P)\mathcal{C}$) [1;2]. For a simple numerical measure of the closeness of $P\mathcal{C}$ to being contained in $Q\mathcal{C}$, $\text{tr}(PQP)$ recommends itself (or, if you like, $n^{-1} \text{tr}(PQP)$). If $Q\mathcal{C}$ is finite-dimensional, one may ask for a modification of Formula 2 which treats $P$ and $Q$ symmetrically.

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

**Formula 4.** $\text{tr}(PQP) = \text{tr}(QPQ) = (GH)^{-1} \sum k_1 G(x_k; y_1) H(y_1; x_k)$.

---

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

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

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
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 \( \{x_1, \ldots, x_n\}, \{x_1^*, \ldots, x_n^*\} \) of \( P\mathcal{C}, \{y_1, \ldots, y_m\}, \{y_1^*, \ldots, y_m^*\} \) of \( Q\mathcal{C} \). It equals

\[
\sum_{kl} (y_l, x_k^*)(x_k, y_l^*) = \sum_{kl} ((x_k, y_l^*)y_l, x_k^*) = \sum_k (Qx_k, x_k^*) = \text{tr}(PQP),
\]

by Formulas 1 and 2.

Formula 5.

\[
c_v(PQP) = c_v(QPQ)
\]

\[
= (GH)^{-1} \sum G(x_{k_1}; y_{l_1}) \cdots (x_{k_{\nu}}; y_{l_{\nu}})H(y_{l_1}; x_{k_1}) \cdots (y_{l_{\nu}}; x_{k_{\nu}}).
\]

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

Proof. The equation

\[
G(x_{k_1}; z_1) \cdots (x_{k_{\nu}}; z_{\nu}) = G_{\nu}(z_1 \otimes \cdots \otimes z_{\nu}, G_{k_1}^* \cdots k_{\nu})
\]

defines an element \( G_{k_1}^* \cdots k_{\nu} \) of \( \mathcal{C}^\nu \), the tensor product of \( \nu \) copies of \( \mathcal{C} \). Extend \( \{x_1, \ldots, x_n\} \) to a basis of \( \mathcal{C} \) by adjoining an orthonormal basis \( \{x_{n+1}, x_{n+2}, \ldots\} \) of \( (1-P)\mathcal{C} \). The elements \( x_{\nu_1} \otimes \cdots \otimes x_{\nu_\nu} \) form a linear basis of \( \mathcal{C}^\nu \). By considering its scalar products with these basis vectors, \( G_{k_1}^* \cdots k_{\nu} \) is identified as

\[
\frac{1}{\nu!} \sum_{l_1 \cdots l_{\nu}} \epsilon_{l_1 \cdots l_{\nu}} x_{l_1}^* \otimes \cdots \otimes x_{l_{\nu}}^* = x_{[k_1} \otimes \cdots \otimes x_{k_{\nu}]}^*.
\]

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

The easily-proved analog of Formula 1 is

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

Formula 3 in the new notation reads

\[
c_v(PAP) = \sum_{k_1 \cdots k_{\nu}} (Ax_{k_1} \otimes \cdots \otimes Ax_{k_{\nu}}, x_{[k_1}^* \otimes \cdots \otimes x_{k_{\nu}]}^*);
\]
this is not disturbed if the subscripts of the $Ax_{k_{i}}$ are also bracketed.

The right side of Formula 5 becomes

$$\sum_{k_{1}\cdots k_{v}, l_{1}\cdots l_{v}} (y_{l_{1}} \otimes \cdots \otimes y_{l_{v}}, x_{[k_{1} \otimes \cdots \otimes x_{k_{v}}]} (x_{k_{1}} \otimes \cdots \otimes x_{k_{v}}, y_{[l_{1} \otimes \cdots \otimes y_{l_{v}}]}).$$

By the analog of Formula 1 this is equal to

$$\sum_{k_{1}\cdots k_{v}} (Qx_{[k_{1} \otimes \cdots \otimes x_{k_{v}}]}, x_{[k_{1} \otimes \cdots \otimes x_{k_{v}}]},$$

and by Formula 3 this is $c_{r}(PQP)$, as claimed.

The analogy to the special case, Formula 4, could be strengthened by mentioning that $Pz_{[i \otimes \cdots \otimes Pz_{v}]} = P_{r}(z_{1} \otimes \cdots \otimes z_{v})$, where $P_{r}$ is the hermitian projection on the subspace of $\mathfrak{X}$ linearly spanned by antisymmetrized products of elements of $P\mathfrak{X}$. The $x_{[k_{1} \otimes \cdots \otimes x_{k_{v}}]}$ and the $x_{[k_{1} \otimes \cdots \otimes x_{k_{v}}]}$ are almost biorthonormal bases of $P_{r}\mathfrak{X}$:

$$(x_{[l_{1} \otimes \cdots \otimes x_{l_{v}}]}, x_{[k_{1} \otimes \cdots \otimes x_{k_{v}}]}^{*}) = \frac{1}{v!} \delta_{k_{1}\cdots k_{v}}.$$