COMPRESSIONS TO FINITE-DIMENSIONAL SUBSPACES

CHANDLER DAVIS

Let $\mathcal{H}$ 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 $P\mathcal{H}$ 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 $k$th 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 $P\mathcal{H}$ 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. $x_1, \ldots, x_n$ 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 $k$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 $P\mathcal{H}$ 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. $\text{tr}(PAP) = G^{-1} \sum_k G(x_k; Ax_k)$.

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

$$(x_i, x_j) = \sum_{\rho\sigma} T_{ip} \bar{T}_{j\sigma}(\xi_\rho, \xi_\sigma) = \sum_\rho T_{ip} \bar{T}_{j\rho} = (TT^*)_{ij}$$
gives $G = |\det T|^2$. \((*)\) denotes conjugate transpose.) It remains to prove that, analogously, \(\sum_k G(x_k; Ax_k) = |\det T|^2 \operatorname{tr}(PAP)\). Now \((Ax_i, x_j) = (PAPx_i, x_j) = \sum_p \lambda_i \lambda_j T_{ip} T^*_{jp}\). 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

$$\sum_k \det R^k = \sum_k \sum_{\rho_1, \ldots, \rho_n} \epsilon_{\rho_1 \cdots \rho_n} T_{\rho_1, \ldots, \rho_n} \cdots T_{n, \rho_n} \lambda_{\rho_k}$$

the summation over $k$ may be carried out first: $\sum_k \lambda_{\rho_k} = \operatorname{tr}(PAP)$ for any $(\rho_1, \ldots, \rho_n)$ 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 $c_\nu$, where for any $B$

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

that is, $c_\nu$ is the $\nu$th elementary symmetric polynomial of the eigenvalues. Extend the notation: $G(x_k; z_1)(x_k; z_2)$ is the determinant of the matrix which differs from the Gramian in having $k_i$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.** $c_\nu(PAP) = G^{-1} \sum G(x_{k_1}; Ax_{k_1}) \cdots (x_{k_\nu}; Ax_{k_\nu})$. (In this equation summation is over all distinct $\nu$-tuples \(\{k_1, \ldots, k_\nu\}\) 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 $Q\mathcal{C}$, \(\operatorname{tr}(PQP)\) recommends itself (or, if you like, $n^{-1} \operatorname{tr}(PQP)$). 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.** $\operatorname{tr}(PQP) = \operatorname{tr}(QPQ) = (GH)^{-1} \sum_k G(x_k; y_1) H(y_1; x_k)$.

\(\nu = n\) shows the equivalence of Theorem 1 of \([4]\) to Weyl's theorem which it generalizes.

One might prefer replacing $PQP$ by $PQP + (1-P)(1-Q)(1-P) = 1 - P - Q + PQ + PQ$, 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 \( \{x_1, \ldots, x_n\}, \{x_1^*, \ldots, x_n^*\} \) of \( P\mathcal{C} \), \( \{y_1, \ldots, y_m\}, \{y_1^*, \ldots, y_m^*\} \) of \( Q\mathcal{C} \). It equals

\[
\sum_{k_l} (y_{l_1} x_k^*)(x_k, y_{l_1}^*) = \sum_{k_l} ((x_k, y_{l_1}^*)y_{l_1}, x_k^*) = \sum_{k} (Qx_k, x_k^*) = \text{tr}(PQP),
\]

by Formulas 1 and 2.

Formula 5.

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

(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

\[
G(x_{k_1}; z_{l_1}) \cdots (x_{k_\nu}; z_{l_\nu}) = G_{\nu'}(z_{l_1} \otimes \cdots \otimes z_{l_\nu}, G_{k_1}^{*} \cdots k_\nu)
\]
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 \( x_{s_1} \otimes \cdots \otimes x_{s_\nu} \) 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

\[
Pz_{[l_1} \otimes \cdots \otimes Pz_{r]} = \sum_{k_1 \cdots k_\nu} (z_{l_1} \otimes \cdots \otimes z_{l_\nu}, x_{[k_1} \otimes \cdots \otimes x_{k_\nu]}^*) x_{k_1} \otimes \cdots \otimes x_{k_\nu}.
\]

Formula 3 in the new notation reads

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

The right side of Formula 5 becomes

$$\sum_{k_1 \cdots k_p, i_1 \cdots i_p} (y_{i_1} \otimes \cdots \otimes y_{i_p}, x_{[k_1} \otimes \cdots \otimes x_{k_p]})(x_{[k_1} \otimes \cdots \otimes x_{k_p]}, y_{[i_1} \otimes \cdots \otimes y_{i_p]}).$$

By the analog of Formula 1 this is equal to

$$\sum_{k_1 \cdots k_p} (Qx_{[k_1} \otimes \cdots \otimes Qx_{k_p]}, x_{[k_1} \otimes \cdots \otimes x_{k_p]}),$$

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 $Pz_{[1} \otimes \cdots \otimes Pz_n] = P_r(z_{1} \otimes \cdots \otimes z_n)$, where $P_r$ is the hermitian projection on the subspace of $\mathcal{H}$ linearly spanned by antisymmetrized products of elements of $PX$. The $x_{[k_1} \otimes \cdots \otimes x_{k_p]}$ and the $x_{[k_1} \otimes \cdots \otimes x_{k_p]}$ are almost biorthonormal bases of $P_r\mathcal{H}$:

$$(x_{[l_1} \otimes \cdots \otimes x_{l_p]}, x_{[k_1} \otimes \cdots \otimes x_{k_p]} = \frac{1}{\nu!} \epsilon_{k_1 \cdots k_p}.$$ 

References