ON SOME PROBLEMS IN CANONICAL CORRELATIONS*

By R. G. LAHA

Indian Statistical Institute, Calcutta

1. Introduction

It is well-known that if $A(:p \times p) = X_1(:p \times n_1)\, X'_1(:n_1 \times p)$ and $B(:p\times p) = X_2\ (:p \times n_2) X'_2(:n_2 \times p)$ are two p-variate sample dispersion matrices based on n_1 and n_2 d.f. respectively which are independently distributed in Wishart's form with the same parameter matrix Σ and if the roots of the determinantal equation $|A-\theta(A+B)|=0$ be arranged in order of magnitude such that $0<\theta_1 \leqslant \theta_2 \leqslant \theta_3 \ldots \leqslant \theta_k < 1$, then the joint distribution of the roots has been obtained by P. L. Hsu (1939) by analytical methods and independently by S. N. Roy (1939) by the geometrical method.

The density function may be written as

$$F(k, \nu_1, \nu_2) = C(k, \nu_1, \nu_2). \prod_{i=1}^{k} \theta_i^{\frac{\nu_1}{2}-1}(1-\theta_i)^{\frac{\nu_2}{2}-1} \prod_{i>j} (\theta_i - \theta_j) \qquad \ldots \quad (1.1)$$

where $$C(k, \nu_1, \nu_2) = \pi^{k/2} \prod_{i=1}^{k} \frac{\Gamma\left(\frac{\nu_1+\nu_2+k+i}{2}-1\right)}{\Gamma\left(\frac{i}{2}\right).\Gamma\left(\frac{\nu_1+i-1}{2}\right).\Gamma\left(\frac{\nu_2+i-1}{2}\right)},$$

$$k = \min(p, n_1);\ \nu_1 = |n_1 - p| + 1 \ \text{ and }\ \nu_2 = n_2 - p + 1 (n_2 > p).$$

It is also observed by Hsu (1939) that the same form gives the joint distribution of Hotelling's canonical correlations (1936) under the hypothesis that the two sets of variates are independent.

Let the variates $x_1, x_2, \ldots, x_p;\ x_{p+1}, \ldots, x_{p+q}$ follow a $(p+q)$-variate normal distribution with the dispersion matrix

$$\begin{array}{cc} & \begin{array}{cc} p & q \end{array} \\ \begin{array}{c} p \\ q \end{array} & \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma'_{12} & \Sigma_{22} \end{pmatrix}. \end{array}$$

---

* Communicated to the 40th Session of the Indian Science Congress held at Lucknow in January, 1953.

If now N stands for the sample size and the corresponding matrix of sum of squares and products in the sample is denoted by

$$\begin{array}{cc} & \begin{array}{cc} p & q \end{array} \\ \begin{array}{c} p \\ q \end{array} & \begin{pmatrix} S_{11} & S_{12} \\ S'_{12} & S_{22} \end{pmatrix} \end{array}$$

then under the hypothesis H ($\Sigma_{12}=O$), the joint density function of the canonical correlations which are defined as the square-roots of the roots of the determinantal equation

$$|S_{12}S_{22}^{-1}\, S'_{12}-\theta S_{11}|=0$$

will follow the same distribution as (1.1) with

$$k=\min(p,q),\ \nu_1=|p-q|+1 \text{ and } \nu_2=N-p-q(N\geqslant p+q).$$

In the next section we shall, first of all, discuss the effect of adding r extra variates to the q-set (where $q\geqslant p$).

2. EFFECT OF ADDITION OF NEW VARIATES

If the variates $x_1, x_2, \ldots, x_p; x_{p+1}, \ldots, x_{p+q}$ form a $(p+q)$-variate normal population with the dispersion matrix

$$\begin{array}{cc} & \begin{array}{cc} p & q \end{array} \\ \begin{array}{c} p \\ q \end{array} & \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma'_{12} & \Sigma_{22} \end{pmatrix} \end{array}$$

then the canonical correlations in the population as defined by Hotelling (1936) will be given by the square-roots of the p roots θ_i $(0<\theta_1<\theta_2<\ldots<\theta_p<1)$ of the determinantal equation

$$|\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}'-\theta\Sigma_{11}|=0$$

where it is assumed that $q\geqslant p$.

Let now r extra variates be added to the q-set, so that now the variates $x_1, x_2, \ldots, x_p;\ x_{p+1}, x_{p+2}, \ldots, x_{p+q};\ x_{p+q+1}, \ldots, x_{p+q+r}$ form a $(p+q+r)$-variate normal population with the dispersion matrix

$$\begin{array}{cc} & \begin{array}{ccc} p & q & r \end{array} \\ \begin{array}{c} p \\ q \\ r \end{array} & \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma'_{12} & \Sigma_{22} & \Sigma_{23} \\ \Sigma'_{13} & \Sigma'_{23} & \Sigma_{33} \end{pmatrix} \end{array}.$$

Let $$\Lambda_{12} = p(\overset{q}{\Sigma_{12}} : \overset{r}{\Sigma_{13}}) \text{ and } \Lambda_{22} = \begin{array}{c c} & \begin{array}{cc} q & r \end{array} \\ \begin{array}{c} q \\ r \end{array} & \begin{pmatrix} \Sigma_{22} & \Sigma_{23} \\ \Sigma'_{23} & \Sigma_{33} \end{pmatrix} \end{array} \quad \ldots \quad (2.1)$$

Then the canonical correlations in the second case are given by the square-roots of the roots $\Phi_i (0 < \Phi_1 \leqslant \Phi_2 \leqslant \ldots \leqslant \Phi_p < 1)$ of the determinantal equation

$$|\Lambda_{12}\Lambda_{22}^{-1}\Lambda'_{12} - \Phi\Sigma_{11}| = 0.$$

We now prove

Theorem 1: *The addition of r extra variates to the q-set can never decrease the sum of the squares of the canonical correlations.*

Proof: From the property of the roots of a determinantal equation it follows evidently that

$$\left.\begin{aligned} \Sigma\theta_i &= \operatorname{tr}(\Sigma_{12}\Sigma_{22}^{-1}\Sigma'_{12}\Sigma_{11}^{-1}), \\ \Sigma\Phi_i &= \operatorname{tr}(\Lambda_{12}\Lambda_{22}^{-1}\Lambda'_{12}\Sigma_{11}^{-1}). \end{aligned}\right\} \quad \ldots \quad (2.2)$$

Hence $$\begin{aligned} \Sigma\Phi_i - \Sigma\theta_i &= \operatorname{tr}(\Lambda_{12}\Lambda_{22}^{-1}\Lambda'_{12}\Sigma_{11}^{-1} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma'_{12}\Sigma_{11}^{-1}) \\ &= \operatorname{tr}[(\Lambda_{12}\Lambda_{22}^{-1}\Lambda'_{12} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma'_{12})\Sigma_{11}^{-1}]. \end{aligned} \quad \ldots \quad (2.3)$$

First of all, we should show that the matrix $(\Lambda_{12}\Lambda_{22}^{-1}\Lambda'_{12} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma'_{12})$ is non-negative definite. For this purpose, let $\Lambda_{22}^{-1} = \begin{array}{c c} & \begin{array}{cc} q & r \end{array} \\ \begin{array}{c} q \\ r \end{array} & \begin{pmatrix} A & B \\ B' & C \end{pmatrix} \end{array}$ where Λ_{22} is defined in (2.1).

Now following the method of Hotelling (1943) for finding the inverse of a matrix we have

$$\begin{array}{c c} & \begin{array}{cc} q & r \end{array} \\ \begin{array}{c} q \\ r \end{array} & \begin{pmatrix} \Sigma_{22} & \Sigma_{23} \\ \Sigma'_{23} & \Sigma_{33} \end{pmatrix} \end{array} \begin{pmatrix} A & B \\ B' & C \end{pmatrix} = \begin{array}{c c} & \begin{array}{cc} q & r \end{array} \\ \begin{array}{c} q \\ r \end{array} & \begin{pmatrix} I & O \\ O & I \end{pmatrix} \end{array} \quad \ldots \quad (2.4)$$

Now on solving the equations for A, B, C in (2.4), we have

$$A = \Sigma_{22}^{-1} + \Sigma_{22}^{-1}\Sigma_{23}Q^{-1}\Sigma'_{23}\Sigma_{22}^{-1},$$

$$B = -\Sigma_{22}^{-1}\Sigma_{23}Q^{-1},$$

and $$C = \Sigma_{33}^{-1} + \Sigma_{33}^{-1}\Sigma'_{23}\Sigma_{22}^{-1}\Sigma_{23}Q^{-1},$$

where $$Q = (\Sigma_{33} - \Sigma'_{23}\Sigma_{22}^{-1}\Sigma_{23}) \quad \ldots \quad (2.5)$$

(assuming, of course, that Q should be non-singular).

It may be noted that $CQ = I$, so that $C = Q^{-1}$.

Hence $\Lambda_{12}\,\Lambda_{22}^{-1}\Lambda'_{12} =$

$$= (\Sigma_{12} : \Sigma_{13})\begin{pmatrix} A & B \\ B' & C \end{pmatrix}\begin{pmatrix} \Sigma'_{12} \\ \Sigma'_{13} \end{pmatrix}$$

$$= \Sigma_{12}\Sigma_{22}^{-1}\Sigma'_{12}+\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{23}Q^{-1}\Sigma'_{23}\Sigma_{22}^{-1}\Sigma'_{12}-$$

$$-\Sigma_{13}Q^{-1}\Sigma'_{23}\Sigma_{22}^{-1}\Sigma'_{12}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{23}Q^{-1}\Sigma'_{13}+\Sigma_{13}\Sigma_{33}^{-1}\Sigma'_{13}+$$

$$+\Sigma_{13}\Sigma_{33}^{-1}\Sigma'_{23}\Sigma_{22}^{-1}\Sigma_{23}Q^{-1}\Sigma'_{13}. \qquad \ldots \quad (2.6)$$

Thus $\Lambda_{12}\Lambda_{22}^{-1}\Lambda'_{12}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma'_{12}$

$$= DQ^{-1}D'-\Sigma_{13}Q^{-1}D'-DQ^{-1}\Sigma'_{13}+\Sigma_{13}\Sigma_{33}^{-1}\Sigma'_{13}+$$

$$+\Sigma_{13}\Sigma_{33}^{-1}\Sigma'_{23}\Sigma_{22}^{-1}\Sigma_{23}Q^{-1}\Sigma'_{13} \qquad \ldots \quad (2.7)$$

where $D = \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{23}$ and Q as defined in (2.5).

But $\Sigma_{13}\,\Sigma_{33}^{-1}\Sigma'_{13}+\Sigma_{13}\Sigma_{33}^{-1}\Sigma'_{23}\Sigma_{22}^{-1}\Sigma_{23}Q^{-1}\Sigma'_{13}$

$$= \Sigma_{13}[\Sigma_{33}^{-1}+\Sigma_{33}^{-1}\Sigma'_{23}\Sigma_{22}^{-1}\Sigma_{23}Q^{-1}]\Sigma'_{13}$$

$$= \Sigma_{13}\Sigma_{33}^{-1}[Q+\Sigma'_{23}\Sigma_{22}^{-1}\Sigma_{23}]Q^{-1}\Sigma'_{13} = \Sigma_{13}Q^{-1}\Sigma'_{13}. \qquad \ldots \quad (2.8)$$

Substituting (2.8) in (2.7), we have at once

$$\Lambda_{12}\Lambda_{22}^{-1}\Lambda'_{12}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma'_{12}$$

$$= (\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{23}-\Sigma_{13})(\Sigma_{33}-\Sigma'_{23}\Sigma_{22}^{-1}\Sigma_{23})^{-1}(\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{23}-\Sigma_{13})'. \quad \ldots \quad (2.9)$$

But the matrix $(\Sigma_{33}-\Sigma'_{23}\Sigma_{22}^{-1}\Sigma_{23})^{-1}$ must be always non-negative definite.

Thus matrix (2.9) is non-negative definite and Σ_{11}^{-1} is non-negative definite. But since the trace of a non-negative definite matrix can never be negative, we must have from (2.3) $\Sigma\Phi_i-\Sigma\theta_i \geqslant 0$ or, $\Sigma\Phi_i \geqslant \Sigma\theta_i$.

Hence the theorem is proved.

Further, since the trace of a non-negative definite matrix is zero when and only when the matrix is a null matrix, it follows :

Corollary 1: *The necessary and sufficient condition for* $\Sigma\phi_i = \Sigma\theta_i$ *is that* $\Lambda_{12}\Lambda_{22}^{-1}\Lambda'_{12} = \Sigma_{12}\Sigma_{22}^{-1}\Sigma'_{12}$.

We also prove

Theorem 2: *The necessary and sufficient condition for the sum of the squares of the canonical correlations to remain the same in both situations (that is, before and after the addition of r extra variates to the q-set) is that all the partial canonical correlations in the population between the p-set and the r-set after eliminating the effect of the q-set, should vanish.*

Proof: The partial canonical correlations in the population between the p-set and the r-set, after eliminating the effect of the q-set are given by the roots of the determinantal equation

$$|(\Sigma_{13}-\Sigma_{12}\Sigma_{22}{}^{-1}\Sigma_{23})(\Sigma_{33}-\Sigma'_{23}\Sigma_{22}{}^{-1}\Sigma_{23})^{-1}(\Sigma_{13}-\Sigma_{12}\Sigma_{22}{}^{-1}\Sigma_{23})'-$$
$$-\psi(\Sigma_{11}-\Sigma_{12}\Sigma_{22}{}^{-1}\Sigma'_{12})|=0 \quad \dots \quad (2.10)$$

as defined by A. C. Das (1948).
Now, from (2.9), (2.10) reduces to the form

$$|(\Lambda_{12}\Lambda_{22}{}^{-1}\Lambda'_{12}-\Sigma_{12}\Sigma_{22}{}^{-1}\Sigma'_{12})-\psi(\Sigma_{11}-\Sigma_{12}\Sigma_{22}{}^{-1}\Sigma'_{12})|=0. \quad \dots \quad (2.11)$$

Hence the necessary and sufficient condition for all ψ_i's to vanish is that

$$\Lambda_{12}\Lambda_{22}{}^{-1}\Lambda'_{12}=\Sigma_{12}\Sigma_{22}{}^{-1}\Sigma'_{12}.$$

The theorem follows then, at once, from the corollary to Theorem 1.

Thus we reach the following interesting conclusion.

Testing the hypothesis as to whether the addition of r extra variates to the q-set does not increase the sum of the squares of the canonical correlations is equivalent to testing that all the population partial canonical correlations between the p-set and the r-set vanish for the fixed q-set.

The hypothesis $H(\Sigma\phi_i=\Sigma\theta_i)$ can be easily tested using the null distribution of the partial canonical correlations as obtained by A. C. Das (1948). Now instead of adding r extra variates to the q-set (where $q \geqslant p$) if we add to the p-set, the number of canonical correlations will not remain the same. After the addition it will be q or $p+r$ according as $q \leqslant p+r$ or $q \geqslant p+r$ respectively.

In the case where $q \geqslant p+r$ the difference in the sums of the squares of the canonical correlations is equal to

$$\mathrm{tr}\{\Lambda'_{21}\Sigma_{22}{}^{-1}\Lambda_{21}\Lambda_{11}{}^{-1}-\Sigma_{12}\Sigma_{22}{}^{-1}\Sigma'_{12}\Sigma_{11}{}^{-1}\} \quad \dots \quad (2.12)$$

where $$\Lambda_{11}=\begin{array}{c|cc} & p & r \\ p & \Sigma_{11} & \Sigma_{13} \\ r & \Sigma_{13}' & \Sigma_{33}\end{array} \quad \text{and} \quad \Lambda_{21}=q(\overset{p}{\Sigma_{12}'}:\overset{r}{\Sigma_{23}}).$$

But in virtue of the relations $\mathrm{tr}\,(AB) = \mathrm{tr}\,(BA)$ and $\mathrm{tr}(A+B) = \mathrm{tr}A+\mathrm{tr}B$, (2.12) reduces to the form

$$\mathrm{tr}[(\Lambda_{21}\Lambda_{11}^{-1}\Lambda'_{21}-\Sigma'_{12}\Sigma_{11}^{-1}\Sigma_{12})\Sigma_{22}^{-1}]. \qquad \dots \quad (2.13)$$

Now proceeding exactly as in the previous method it can be proved easily that Theorems 1 and 2 also hold for both the cases: (i) $q \leqslant p+r$, (ii) $q \geqslant p+r$.

REFERENCES

DAS, A. C. (1948): On the canonical multiple and partial correlations. *Proc. Ind. Sc. Cong.*, Patna.

HOTELLING, H. (1936): Relations between two sets of variates. *Biometrika*, 28, 328–377.

——— (1943): Some new methods in matrix calculation. *Ann. Math. Stat.*, 14, 1–34.

HSU, P. L. (1939): On the distribution of the roots of certain determinantal equations. *Ann. Eugen.*, 9, 250–258.

ROY, S. N. (1939): *p*-statistics or some generalisations in the analysis of variance appropriate to multivariate problems. *Sankhyā*, 4, 381–396.

Paper received: October, 1952.