CANONICAL CORRELATION AND CANONICAL VARIATE

Suppose we wish to determine the degree of linear relationship between a single variable, Y and a linear combination of the variables in a given set say $X_1, \cdots, X_p$, then the multiple correlation coefficient may be an appropriate statistic to use. Further, if we have two sets, a set of $Y's$ say $Y_1, \cdots, Y_q$ as well as a set of $X's$, and that we wish to determine the degree of linear relationship between linear combinations of variables in the two sets, an appropriate statistical procedure in this case is known as canonical correlation analysis. One may view the method as an extension of multiple correlation analysis; however, it might be more appropriate to view multiple correlation as a special case of canonical correlation. Each of these linear combinations is called a canonical variate and the correlation between them is called canonical correlation. For example, i) one set of variables may be measurements of physical characteristics such as various lengths and breadths of skulls, the other variables may be measurements of mental characteristics, such as scores on intelligence tests, ii) one set may be marks obtained by a group of students in art subject and the other set may marks obtained by the same group of students in science subjects.

Procedure

If the two sets of variables are very large then the study between the two sets of variable will be made easier by considering the linear combinations of the two sets of variable instead of considering them as they are. Consider only those linear combinations of the two sets of variables that has the maximum correlation subject to certain conditions, and then determine the pair of linear combination having the maximum correlation among such pair of linear combination uncorrelated with the initially selected pair. This process continues until all the linear combinations are completely specified.

Let $\underline{X}$ be a random vector of $p-$components with $E\,\underline{X} = \underline{0}$ and covariance matrix Σ (which is assumed to be positive definite). We partition $\underline{x}$ into sub vectors of p_1 and p_2 components respectively.

$$\underline{X} = \begin{pmatrix} \underline{X}^{(1)} \\ \hline \underline{X}^{(2)} \end{pmatrix}, \text{ with } p_1 + p_2 = p, \text{ and } p_1 \le p_2 .$$

The covariance matrix is partitioned similarly into p_1 and p_2 rows and columns

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Let us consider an arbitrary linear combination, $U = \underline{\alpha}'\,\underline{X}^{(1)}$ of the components of $\underline{X}^{(1)}$ and an arbitrary linear function $V = \underline{\gamma}'\,\underline{X}^{(2)}$ of the components of $\underline{X}^{(2)}$.

where $\underline{\alpha}$ and $\underline{\gamma}$ are to be so chosen that the correlation between U and V is the maximum. Since the correlation between two random variables is invariant with respect to change of scale, the vector $\underline{\alpha}$ and $\underline{\gamma}$ can be normalized arbitrarily. So, we may choose $\underline{\alpha}$ and $\underline{\gamma}$ such that

$$\text{Var}(U) = \text{Var}(V) = 1 \tag{9.1}$$

i.e. $\text{Var}(U) = EU^2 = E\,\underline{\alpha}'\,\underline{X}^{(1)}\,\underline{X}^{(1)'}\,\underline{\alpha} = \underline{\alpha}'\Sigma_{11}\underline{\alpha}$, and

$$\text{Var}(V) = EV^2 = E\underline{\gamma}'\,\underline{X}^{(2)}\,\underline{X}^{(2)'}\,\underline{\gamma} = \underline{\gamma}'\Sigma_{22}\,\underline{\gamma} \tag{9.2}$$

because, $EU = E\,\underline{\alpha}'\,\underline{X}^{(1)} = \underline{\alpha}'E\,\underline{X}^{(1)} = 0$, and $EV = 0$. Then

$$\text{Corr}(U,V) = \text{Cov}(U,V) = EUV = E\,\underline{\alpha}'X^{(1)}\,\underline{X}^{(2)'}\,\underline{\gamma} = \underline{\alpha}'\Sigma_{12}\underline{\gamma} \tag{9.3}$$

Thus the problem is to find $\underline{\alpha}$ and $\underline{\gamma}$ to maximize (9.3) subject to condition (9.1) and (9.2).

Let

$$\phi = \underline{\alpha}'\,\Sigma_{12}\,\underline{\gamma} - \frac{1}{2}\lambda(\underline{\alpha}'\,\Sigma_{11}\underline{\alpha} - 1) - \frac{1}{2}\mu(\underline{\gamma}'\Sigma_{22}\underline{\gamma} - 1)$$

where λ and μ are Lagrange multipliers. Taking the derivatives of ϕ with respect to the components of $\underline{\alpha}$ and $\underline{\gamma}$, and equating them to zero.

$$\frac{\partial\phi}{\partial\underline{\alpha}} = \Sigma_{12}\,\underline{\gamma} - \lambda\Sigma_{11}\underline{\alpha} = \underline{0} \tag{9.4}$$

$$\frac{\partial\phi}{\partial\underline{\gamma}} = \Sigma_{12}'\,\underline{\alpha} - \mu\Sigma_{22}\underline{\gamma} = \underline{0} \tag{9.5}$$

Multiply (9.4) on the left by $\underline{\alpha}'$ and by (9.5) on the left by $\underline{\gamma}'$, we get

$$\underline{\alpha}'\,\Sigma_{12}\,\underline{\gamma} - \lambda\underline{\alpha}'\,\Sigma_{11}\underline{\alpha} = \underline{0}, \text{ and } \quad \underline{\gamma}'\,\Sigma_{12}'\,\underline{\alpha} - \mu\underline{\gamma}'\,\Sigma_{22}\underline{\gamma} = \underline{0}$$

Since $\underline{\alpha}'\,\Sigma_{11}\,\underline{\alpha} = 1$, and $\underline{\gamma}'\,\Sigma_{22}\,\underline{\gamma} = 1$

$$\Rightarrow\ \lambda = \underline{\alpha}'\Sigma_{12}\underline{\gamma}, \text{ and } \quad \mu = \underline{\gamma}'\Sigma_{12}\underline{\alpha}$$

$$\Rightarrow\ \lambda = \mu = \underline{\alpha}'\Sigma_{12}\underline{\gamma} \tag{9.6}$$

Thus, the normal equations (9.4) and (9.5) can be written as

$$-\lambda\Sigma_{11}\underline{\alpha} + \Sigma_{12}\,\underline{\gamma} = \underline{0} \tag{9.7}$$

$$\text{and } \Sigma_{12}'\,\underline{\alpha} - \lambda\Sigma_{22}\underline{\gamma} = \underline{0} \tag{9.8}$$

$$\text{or } \begin{pmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{pmatrix}\begin{pmatrix} \underline{\alpha} \\ \underline{\gamma} \end{pmatrix} = \underline{0} \tag{9.9}$$

Since $\begin{pmatrix} \underline{\alpha} \\ \underline{\gamma} \end{pmatrix} \neq \underline{0}$, $\quad \Rightarrow\ \underline{\alpha} \neq \underline{0}$, $\underline{\gamma} \neq \underline{0}$. we must have

$$\begin{vmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{vmatrix} = 0 \tag{9.10}$$

i.e. $\begin{pmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{pmatrix}$ must be a singular matrix.

The function $\begin{vmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{vmatrix}$ is a polynomial in λ of degree p. Therefore the equation (9.10) will give p solutions in λ say $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$.

From equation (9.6), we observe that $\lambda = \underline{\alpha}'\Sigma_{12}\underline{\gamma}$ is the correlation between $U = \underline{\alpha}'\underline{X}^{(1)}$ and $V = \underline{\gamma}'\underline{X}^{(2)}$, when $\underline{\alpha}$ and $\underline{\gamma}$ must satisfy (9.9) for some values of λ.

Since we want the maximum correlation, we take $\lambda = \lambda_1$,

Let $\underline{\alpha}^{(1)}$ and $\underline{\gamma}^{(1)}$ be the normalized solution of

$$\begin{pmatrix} -\lambda_1\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda_1\Sigma_{22} \end{pmatrix}\begin{pmatrix} \underline{\alpha} \\ \underline{\gamma} \end{pmatrix} = \underline{0}\text{, then}$$

$U_1 = \underline{\alpha}^{(1)'}\underline{X}^{(1)}$ and $V_1 = \underline{\gamma}^{(1)'}\underline{X}^{(2)}$ are normalized linear combinations of $\underline{X}^{(1)}$ and $\underline{X}^{(2)}$ respectively, with maximum correlation. U_1 and V_1 are called the first set of canonical variates.

We now consider finding the second set of linear combination of $\underline{X}^{(1)}$, say $U = \underline{\alpha}'\underline{X}^{(1)}$ and of $\underline{X}^{(2)}$ say $V = \underline{\gamma}'\underline{X}^{(2)}$, such that of all linear combinations uncorrelated with U_1 and V_1; these have maximum correlation. Non correlation with U_1 and V_1 is same as

$$0 = \mathrm{Cov}(U, U_1) = E(\underline{\alpha}'\underline{x}^{(1)}\underline{x}^{(1)'}\underline{\alpha}^{(1)}) = \underline{\alpha}'\Sigma_{11}\underline{\alpha}^{(1)}$$

$$0 = \mathrm{Cov}(V, V_1) = E(\underline{\gamma}'\underline{x}^{(2)}\underline{x}^{(2)'}\underline{\gamma}^{(1)}) = \underline{\gamma}'\Sigma_{22}\underline{\gamma}^{(1)}$$

$$0 = \mathrm{Cov}(U, V_1) = E(\underline{\alpha}'\underline{x}^{(1)}\underline{x}^{(2)'}\underline{\gamma}^{(1)}) = \underline{\alpha}'\Sigma_{12}\underline{\gamma}^{(1)}$$

$$0 = \mathrm{Cov}(V, U_1) = E(\underline{\gamma}'\underline{x}^{(2)}\underline{x}^{(1)'}\underline{\alpha}^{(1)}) = \underline{\gamma}'\Sigma_{21}\underline{\alpha}^{(1)}$$

Let

$$\phi_2 = \underline{\alpha}'\Sigma_{12}\underline{\gamma} - \frac{1}{2}\lambda(\underline{\alpha}'\Sigma_{11}\underline{\alpha} - 1) - \frac{1}{2}\mu(\underline{\gamma}'\Sigma_{22}\underline{\gamma} - 1) - \upsilon_1\underline{\alpha}'\Sigma_{11}\underline{\alpha}^{(1)} - \upsilon_2\underline{\gamma}'\Sigma_{22}\underline{\gamma}^{(1)}$$

Taking the derivatives of ϕ_2 with respect to the components of $\underline{\alpha}$ and $\underline{\gamma}$, and equating them to zero, gives

$$\frac{\partial\phi_2}{\partial\underline{\alpha}} = \underline{0} = \Sigma_{12}\underline{\gamma} - \lambda\Sigma_{11}\underline{\alpha} - \upsilon_1\Sigma_{11}\underline{\alpha}^{(1)} \tag{9.11}$$

$$\frac{\partial\phi_2}{\partial\underline{\gamma}} = \underline{0} = \Sigma_{12}'\underline{\alpha} - \mu\Sigma_{22}\underline{\gamma} - \upsilon_2\Sigma_{22}\underline{\gamma}^{(1)} \tag{9.12}$$

Multiply (9.11) on the left by $\underline{\alpha}^{(1)'}$ and by (9.12) on the left by $\underline{\gamma}^{(1)'}$, we get

$$\underline{\alpha}^{(1)'}\Sigma_{12}\underline{\gamma} - \lambda\underline{\alpha}^{(1)'}\Sigma_{11}\underline{\alpha} - \upsilon_1\underline{\alpha}^{(1)'}\Sigma_{11}\underline{\alpha}^{(1)} = 0$$

and $\underline{\gamma}^{(1)'} \Sigma_{12}' \underline{\alpha} - \mu \underline{\gamma}^{(1)'} \Sigma_{22} \underline{\gamma} - \upsilon_2 \underline{\gamma}^{(1)'} \Sigma_{22} \underline{\gamma}^{(1)} = 0$

$\Rightarrow \; -\upsilon_1 \underline{\alpha}^{(1)'} \Sigma_{11} \underline{\alpha}^{(1)} = 0$, and $\quad -\upsilon_2 \underline{\gamma}^{(1)'} \Sigma_{22} \underline{\gamma}^{(1)} = 0$

$\Rightarrow \; \upsilon_1 = \upsilon_2 = 0$, because $\mathrm{V}(U_1) = \underline{\alpha}^{(1)'} \Sigma_{11} \underline{\alpha}^{(1)} = 1 = \mathrm{V}(V_1)$

Thus the normal equations are just the same as (9.9)

So the procedure to find the second set of linear combination is to solve equation (9.9) for $\lambda = \lambda_2$, and call the vectors as $\underline{\alpha}^{(2)}$, $\underline{\gamma}^{(2)}$ and the corresponding normalized linear combinations are

$U_2 = \underline{\alpha}^{(2)'} \underline{x}^{(1)}$ and $V_2 = \underline{\gamma}^{(2)'} \underline{x}^{(2)}$ of $\underline{x}^{(1)}$ and $\underline{x}^{(2)}$ respectively, with maximum correlation. U_2 and V_2 are called the second set of canonical variates.

Let us have the sets of canonical variates

$$
\begin{array}{lll}
U_1 = \underline{\alpha}^{(1)'} \underline{x}^{(1)} & V_1 = \underline{\gamma}^{(1)'} \underline{x}^{(2)} & \mathrm{Cov}(U_1, V_1) = \underline{\alpha}^{(1)'} \Sigma_{12} \underline{\gamma}^{(1)} \\
U_2 = \underline{\alpha}^{(2)'} \underline{x}^{(1)} & V_2 = \underline{\gamma}^{(2)'} \underline{x}^{(2)} & \mathrm{Cov}(U_2, V_2) = \underline{\alpha}^{(2)'} \Sigma_{12} \underline{\gamma}^{(2)} \\
\vdots & \vdots & \vdots \\
U_r = \underline{\alpha}^{(r)'} \underline{x}^{(1)} & V_r = \underline{\gamma}^{(r)'} \underline{x}^{(2)} & \mathrm{Cov}(U_r, V_r) = \underline{\alpha}^{(r)'} \Sigma_{12} \underline{\gamma}^{(2)}
\end{array}
$$

We have to determine

$U = \underline{\alpha}' \underline{x}^{(1)}$, and $V = \underline{\alpha}' \underline{x}^{(2)}$, such that of all linear combinations uncorrelated with $U_1, V_1, \cdots, U_r, V_r$ have maximum correlation.

The condition that U be uncorrelated with U_i is

$$\mathrm{Cov}(U, U_i) = 0, \quad i = 1, 2, \cdots, r$$

$$= E(\underline{\alpha}' \underline{x}^{(1)} \underline{x}^{(1)'} \underline{\alpha}^{(i)}) = \underline{\alpha}' \Sigma_{11} \underline{\alpha}^{(i)} \tag{9.13}$$

If $\lambda_i \neq 0$, from equation (9.7), we get

$$\Sigma_{11} \underline{\alpha}^{(i)} = \frac{1}{\lambda_i} \Sigma_{12} \underline{\gamma}^{(i)}, \text{ and therefore, } \; \underline{\alpha}' \Sigma_{11} \underline{\alpha}^{(i)} = \frac{1}{\lambda_i} \underline{\alpha}' \Sigma_{12} \underline{\gamma}^{(i)}$$

$$\mathrm{Cov}(U, U_i) = \frac{1}{\lambda_i} \mathrm{Cov}(U, V_i) = 0$$

The condition that V be uncorrelated with V_i is

$$\mathrm{Cov}(V, V_i) = 0, \quad i = 1, 2, \cdots, r$$

$$= E(\underline{\gamma}' \underline{x}^{(2)} \underline{x}^{(2)'} \underline{\gamma}^{(i)}) = \underline{\gamma}' \Sigma_{22} \underline{\gamma}^{(i)} \tag{9.14}$$

From equation (9.8), we get

$$\Sigma_{22}\underline{\gamma}^{(i)} = \frac{1}{\lambda_i}\Sigma_{21}\underline{\alpha}^{(i)} \qquad \text{or} \qquad \underline{\gamma}\Sigma_{22}\underline{\gamma}^{(i)} = \frac{1}{\lambda_i}\underline{\gamma}\,\Sigma_{21}\underline{\alpha}^{(i)}$$

$$\Rightarrow \operatorname{Cov}(V, V_i) = \frac{1}{\lambda_i}\operatorname{Cov}(V, U_i) = 0$$

We now maximize $EU_{r+1}V_{r+1}$, choosing $\underline{\alpha}$ and $\underline{\gamma}$ to satisfy equations (9.1), (9.2) and (9.13), (9.14) for $i = 1, 2, \cdots, r$.

Consider

$$\phi_{r+1} = \underline{\alpha}'\Sigma_{12}\,\underline{\gamma} - \frac{1}{2}\lambda(\underline{\alpha}'\Sigma_{11}\underline{\alpha} - 1) - \frac{1}{2}\mu(\underline{\gamma}'\Sigma_{22}\underline{\gamma} - 1) + \sum_{i=1}^{r}\upsilon_i\,\underline{\alpha}'\Sigma_{11}\underline{\alpha}^{(i)} + \sum_{i=1}^{r}\theta_i\,\underline{\gamma}'\Sigma_{22}\underline{\gamma}^{(i)}$$

where λ, $\mu, \upsilon_1, \upsilon_2, \cdots, \upsilon_r$, $\theta_1, \theta_2, \cdots, \theta_r$ are Lagrange multipliers. The vectors of partial derivatives of ϕ_{r+1} with respect to the components of $\underline{\alpha}$ and $\underline{\gamma}$, and set equal to zero, giving

$$\frac{\partial\phi_{r+1}}{\partial\underline{\alpha}} = \Sigma_{12}\,\underline{\gamma} - \lambda\Sigma_{11}\underline{\alpha} + \sum_{i=1}^{r}\upsilon_i\,\Sigma_{11}\underline{\alpha}^{(i)} = \underline{0} \tag{a}$$

$$\frac{\partial\phi_{r+1}}{\partial\underline{\gamma}} = \Sigma_{12}'\,\underline{\alpha} - \mu\Sigma_{22}\underline{\gamma} + \sum_{i=1}^{r}\theta_i\,\Sigma_{22}\underline{\gamma}^{(i)} = \underline{0} \tag{b}$$

Multiplication of (a) on the left by $\underline{\alpha}^{(j)'}$ and (b) on the left by $\underline{\gamma}^{(j)'}$, where $1 \le j \le r$, gives

$$\underline{\alpha}^{(j)'}\Sigma_{12}\,\underline{\gamma} - \lambda\,\underline{\alpha}^{(j)'}\Sigma_{11}\underline{\alpha} + \underline{\alpha}^{(j)'}\sum_{i=1}^{r}\upsilon_i\,\Sigma_{11}\underline{\alpha}^{(i)} = 0$$

and $\underline{\gamma}^{(j)'}\Sigma_{12}'\,\underline{\alpha} - \mu\,\underline{\gamma}^{(j)'}\Sigma_{22}\underline{\gamma} + \underline{\gamma}^{(j)'}\sum_{i=1}^{r}\theta_i\,\Sigma_{22}\underline{\gamma}^{(i)} = 0$

$\Rightarrow$ $\underline{\alpha}^{(j)'}\sum_{i=1}^{r}\upsilon_i\,\Sigma_{11}\underline{\alpha}^{(i)} = 0$, and $\underline{\gamma}^{(j)'}\sum_{i=1}^{r}\theta_i\,\Sigma_{22}\underline{\gamma}^{(i)} = 0$

or $\upsilon_j\,\underline{\alpha}^{(j)'}\Sigma_{11}\underline{\alpha}^{(j)} + \underline{\alpha}^{(j)'}\sum_{i=1}^{r}\upsilon_i\,\Sigma_{11}\underline{\alpha}^{(i)} = 0$, $i \ne j$

and $\theta_j\,\underline{\gamma}^{(j)'}\Sigma_{22}\underline{\gamma}^{(j)} + \underline{\gamma}^{(j)'}\sum_{i=1}^{r}\theta_i\,\Sigma_{22}\underline{\gamma}^{(i)} = 0$, $i \ne j$

$\Rightarrow$ $\upsilon_j\,\underline{\alpha}^{(j)'}\Sigma_{11}\underline{\alpha}^{(j)} = \upsilon_j = 0$, and $\theta_j\,\underline{\gamma}^{(j)'}\Sigma_{22}\underline{\gamma}^{(j)} = \theta_j = 0$

$\Rightarrow$ $\upsilon_j = \theta_j = 0$, $1 \le j \le r$.

Thus the normal equations (a) and (b) are simply (9.7) and (9.8) or alternatively (9.9). So the procedure is as follows:

Solve the polynomial equation (9.10) to get $\lambda_1 > \lambda_2 > \cdots > \lambda_p$, solve the system of equation (9.9) by substituting the roots $\lambda_1, \lambda, \cdots, \lambda_p$ one by one to get the vectors as

$(\underline{\alpha}^{(1)}, \underline{\gamma}^{(1)}),\ (\underline{\alpha}^{(2)}, \underline{\gamma}^{(2)}),\cdots$

which gives the $r-$th pair of canonical variates as

$U_r = \underline{\alpha}^{(r)'}\,\underline{x}^{(1)}$ and $V_r = \underline{\gamma}^{(r)'}\,\underline{x}^{(2)}$ with their correlation as the canonical correlations.

Result: The square of canonical correlation $(\lambda^2{}'s)$ are the characteristic roots of the matrix $(\Sigma_{11}^{-1}\,\Sigma_{12}\,\Sigma_{22}^{-1}\,\Sigma_{21})$.

Proof: We have the normal equations,

$$-\lambda\,\Sigma_{11}\underline{\alpha} + \Sigma_{12}\,\underline{\gamma} = \underline{0} \tag{9.14}$$

$$\Sigma_{12}{}'\,\underline{\alpha} - \lambda\,\Sigma_{22}\,\underline{\gamma} = \underline{0} \tag{9.15}$$

Multiply the equation (9.14) by λ and (9.15) by Σ_{22}^{-1}, we get

$$-\lambda^2\,\Sigma_{11}\underline{\alpha} + \lambda\,\Sigma_{12}\,\underline{\gamma} = \underline{0}$$

and $\Sigma_{22}^{-1}\Sigma_{12}{}'\,\underline{\alpha} - \lambda\,\Sigma_{22}^{-1}\Sigma_{22}\,\underline{\gamma} = \underline{0}$

$$\Rightarrow\ \lambda\,\Sigma_{12}\,\underline{\gamma} = \lambda^2\,\Sigma_{11}\underline{\alpha} \tag{9.16}$$

$$\text{and}\quad \Sigma_{22}^{-1}\Sigma_{12}{}'\,\underline{\alpha} = \lambda\,\underline{\gamma} \tag{9.17}$$

Substituting $\lambda\,\underline{\gamma}$ from equation (9.17) in (9.16)

$$\Sigma_{12}\,(\Sigma_{22}^{-1}\Sigma_{21}\underline{\alpha}) = \lambda^2\,\Sigma_{11}\underline{\alpha}$$

or $\Sigma_{11}^{-1}\Sigma_{12}\,\Sigma_{22}^{-1}\Sigma_{21}\underline{\alpha} = \lambda^2\,\underline{\alpha}$, or $\left|\Sigma_{11}^{-1}\Sigma_{12}\,\Sigma_{22}^{-1}\Sigma_{21} - \lambda^2 I\right|\underline{\alpha} = \underline{0}$

For its solvability θ should be the characteristic roots of $(\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$.

Result: The square of canonical correlations are invariant under nonsingular linear transformation.

Proof: Let $\underline{x} = \begin{pmatrix} \underline{x}^{(1)} \\ \underline{x}^{(2)} \end{pmatrix}$, and $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$. Consider the transformation

$$\underline{y}^{(1)} = A_1\,\underline{x}^{(1)},\ |\,A_1\,| \neq \underline{0},\ \underline{y}^{(2)} = A_2\,\underline{x}^{(2)},\ |\,A_2\,| \neq \underline{0}$$

$$\Rightarrow\ \underline{y} = \begin{pmatrix} \underline{y}^{(1)} \\ \underline{y}^{(2)} \end{pmatrix},\ \Sigma_y = \begin{pmatrix} A_1\Sigma_{11}A_1' & A_1\Sigma_{12}A_2' \\ A_2\Sigma_{21}A_1' & A_2\Sigma_{22}A_2' \end{pmatrix}$$

We have seen that the square of the canonical correlations are the characteristic roots of

$$\left|\Sigma_{11}^{-1}\Sigma_{12}\,\Sigma_{22}^{-1}\Sigma_{21} - \theta I\right| = 0 \tag{9.18}$$

We will show that the roots of the equation (9.18) are same as that of

$$\left| \Sigma_{11}^{-1}(y)\Sigma_{12}(y)\Sigma_{22}^{-1}(y)\Sigma_{21}(y) - \upsilon I \right| = 0$$

or $\left| A_1'^{-1}\Sigma_{11}^{-1}A_1^{-1}\, A_1\Sigma_{12}A_2'\; A_2'^{-1}\Sigma_{22}^{-1}A_2^{-1}\, A_2\Sigma_{21}A_1' - \upsilon I \right| = 0$

or $\left| A_1'^{-1}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}A_1' - \upsilon A_1'^{-1}A_1' \right| = 0$ or $\left| A_1'^{-1}(\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - \upsilon\, I)\, A_1' \right| = 0$

or $\left| A_1'^{-1} \right| \left| A_1' \right| \left| \Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - \upsilon I \right| = 0$

or $\left| \Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - \upsilon I \right| = 0$ (9.19)

Obviously the roots of (9.18) and (9.19) are same.

Result: Multiple correlation is a special case of canonical correlation.

Proof: For multiple correlation we partitioned

$$\underline{x} = \begin{pmatrix} x^{(1)} \\ \underline{x}^{(2)} \end{pmatrix}, \text{ and } \quad \Sigma = \begin{pmatrix} \sigma_{11} & \underline{\sigma}_{12}' \\ \underline{\sigma}_{12} & \Sigma_{22} \end{pmatrix}$$

For such a partition the square of the canonical correlation is the root of the equation

$\left| \sigma_{11}^{-1}\underline{\sigma}_{12}'\,\Sigma_{22}^{-1}\underline{\sigma}_{12} - \theta I_{1\times 1} \right| = 0$, the roots of this equation is $\theta = \dfrac{\underline{\sigma}_{12}'\Sigma_{22}^{-1}\underline{\sigma}_{12}}{\sigma_{11}} = \rho_{1(2\cdots p)}^2$.

Note: If $p_1 = p_2 = 1$, the canonical correlation reduces to the ordinary correlation coefficient.

Characteristic roots: The characteristic roots of a square matrix B are defined as the roots of the characteristic equation $\left| B - \lambda I \right| = 0$.

For example, if $B = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}$, then $\left| B - \lambda I \right| = \begin{vmatrix} 5-\lambda & 2 \\ 2 & 5-\lambda \end{vmatrix} = \lambda^2 - 10\lambda + 21$. The degree of the characteristic equation is the order of the matrix B and the constant term is $\left| B \right|$.