THE UNIVERSITY OF CHICAGO
Graduate School of Business
Business 41912, Spring Quarter 2008, Mr. Ruey S. Tsay

Lecture 8: Canonical Correlation Analysis

Suppose that $\boldsymbol{X}_1 = (X_{11}, \ldots, X_{1p})'$ be a p-dimensional random vector and $\boldsymbol{X}_2 = (X_{21}, \ldots, X_{2q})'$ be a q-dimensional random vector. Without loss of generality, we assume that $p \le q$.
Assume that $E(\boldsymbol{X}_i) = \boldsymbol{\mu}_i$ and $\mathrm{Cov}(\boldsymbol{X}_i) = \boldsymbol{\Sigma}_{ii}$, which are positive definite. In addition, $\mathrm{Cov}(\boldsymbol{X}_1, \boldsymbol{X}_2) = \boldsymbol{\Sigma}_{12}$.
Canonical correlation analysis focuses on the correlation between a linear combination of $\boldsymbol{X}_1$ and a linear combination of $\boldsymbol{X}_2$. The idea is to find a pair of linear combinations that have the largest correlation. Next, we seek the pair of linear combinations having the largest correlation among all pairs that are uncorrelated with the initially selected pair, and so on. The pairs of linear combinations are called canonical variables and their correlations are called canonical correlations.

1 Canonical variates and canonical correlations

Let $U = \boldsymbol{a}'\boldsymbol{X}_1$ be a linear combination of $\boldsymbol{X}_1$ and $V = \boldsymbol{b}'\boldsymbol{X}_2$ be a linear combination of $\boldsymbol{X}_2$. We have $\mathrm{Var}(U) = \boldsymbol{a}'\boldsymbol{\Sigma}_{11}\boldsymbol{a}$, $\mathrm{Var}(V) = \boldsymbol{b}'\boldsymbol{\Sigma}_{22}\boldsymbol{b}$, and $\mathrm{Cov}(U, V) = \boldsymbol{a}'\boldsymbol{\Sigma}_{12}\boldsymbol{b}$. The correlation between U and V is

$$\mathrm{Corr}(U, V) = \frac{\boldsymbol{a}'\boldsymbol{\Sigma}_{12}\boldsymbol{b}}{\sqrt{\boldsymbol{a}'\boldsymbol{\Sigma}_{11}\boldsymbol{a}}\sqrt{\boldsymbol{b}'\boldsymbol{\Sigma}_{22}\boldsymbol{b}}}. \quad (1)$$

Definition:
The *first canonical variate pair* is the pair of linear combinations U_1 and V_1 having unit variances, which maximize the correlation (1).
The *second cononical variate pair* is the pair of linear combinations U_2 and V_2, which maximize the correlation (1) among all choices that are uncorrelated with the first pair of canonical variables.
The definition generalizes to other pairs of canonical variables.

Result 10.1. Assume $p \le q$. Let $\boldsymbol{X} = (\boldsymbol{X}_1', \boldsymbol{X}_2')'$ such that

$$\mathrm{Cov}(()X) = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

such that $\boldsymbol{\Sigma}_{ii}$ are positive definite and of dimension $p \times p$ and $q \times q$, respectively. For coefficient vectors $\boldsymbol{a}_{p\times 1}$ and $\boldsymbol{b}_{q\times 1}$, form the linear combinations $U = \boldsymbol{a}'\boldsymbol{X}_1$ and $V = \boldsymbol{b}'\boldsymbol{X}_2$. Then,

$$\max_{\boldsymbol{a},\boldsymbol{b}} \mathrm{Cor}(U, V) = \rho_1^*$$

is attained by the linear combinations

$$U_1 = \boldsymbol{e}_1'\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{X}_1, \quad V_1 = \boldsymbol{f}_1'\boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{X}_2.$$

The kth pair of canonical variates, $k = 2, 3, \ldots, p$,

$$U_k = \boldsymbol{e}_k'\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{X}_1, \quad V_k = \boldsymbol{f}_k'\boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{X}_2$$

maximizes

$$\text{Cor}(U_k, V_k) = \rho_k^*$$

among those linear combinations uncorrelated with the preceeding $1, \ldots, k-1$ canonical variates. Here $\rho_1^{*2} \geq \rho_2^{*2} \geq \cdots \geq \rho_p^{*2}$ are the eigenvalues of $\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1/2}$ and $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_p$ are the associated eigenvectors. [The quantities $\rho_1^{*2} \geq \rho_2^{*2} \geq \cdots \geq \rho_p^{*2}$ are also the p largest eigenvalues of the matrix $\boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1/2}$ with eigenvectors $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_p$.] Each $\boldsymbol{f}_i$ is proportional to $\boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{e}_i$
In addition, the canonical variates have the following propertities:

1. $\text{Var}(U_k) = \text{Var}(V_k) = 1$
2. $\text{Cov}(U_k, U_\ell) = \text{corr}(U_k, U_\ell) = 0, \quad k \neq \ell$
3. $\text{Cov}(V_k, V_\ell) = 0, \quad k \neq \ell$
4. $\text{Cov}(U_k, V_\ell) = \text{corr}(U_k, V_\ell) = 0$

for $k, \ell = 1, 2, \ldots, p$.

Proof: (Make use of the Cauchy-Schwarz inequality). Let $\boldsymbol{\Sigma}_{ii}^{1/2}$ be the symmetric square root matrix of $\boldsymbol{\Sigma}_{ii}$ for $i = 1$ and 2. Also, let $\boldsymbol{c} = \boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{a}$ and $\boldsymbol{d} = \boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{b}$. Then,

$$\text{Corr}(\boldsymbol{a}'\boldsymbol{X}_1, \boldsymbol{b}'\boldsymbol{X}_2) = \frac{\boldsymbol{a}'\boldsymbol{\Sigma}_{12}\boldsymbol{b}}{\sqrt{\boldsymbol{a}'\boldsymbol{\Sigma}_{11}\boldsymbol{a}}\sqrt{\boldsymbol{b}'\boldsymbol{\Sigma}_{22}\boldsymbol{b}}} = \frac{\boldsymbol{c}'\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{d}}{\sqrt{\boldsymbol{c}'\boldsymbol{c}}\sqrt{\boldsymbol{d}'\boldsymbol{d}}}.$$

By the Cauchy-Schwars inequality,

$$\boldsymbol{c}'\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{d} \leq (\boldsymbol{c}'\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{c})^{1/2}(\boldsymbol{d}'\boldsymbol{d})^{1/2}.$$

Since $\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1/2}$ is a $p \times p$ symmetric matrix, the maximization result of Chapter 2 yields

$$\boldsymbol{c}'\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{c} \leq \lambda_1\boldsymbol{c}'\boldsymbol{c},$$

where λ_1 is the largest eigenvalue of $\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1/2}$. Equality occurs in the prior equation when $\boldsymbol{c} = \boldsymbol{e}_1$, a normalized eigenvector associated with λ_1. The equality of the Cauchy-Schwarz inequality holds if $\boldsymbol{d}$ is proportitional to $\boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{e}_1$. Consequently,

$$\max_{\boldsymbol{a},\boldsymbol{b}} \text{Corr}(\boldsymbol{a}'\boldsymbol{X}_1, \boldsymbol{b}'\boldsymbol{X}_2) = \sqrt{\lambda_1}$$

with equality occurring at $\boldsymbol{a} = \boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{c} = \boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{e}_1$ and $\boldsymbol{b}$ being proportitional to $\boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{e}_1$, where the sign is selected to give positive correlation. We take $\boldsymbol{b} = \boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{f}$. This last correspondence follows by mutliplying both sides of

$$(\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1/2})\boldsymbol{e}_1 = \lambda_1\boldsymbol{e}_1$$

by $\boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1/2}$, yielding

$$\boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1/2}(\boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1/2})\boldsymbol{e}_1 = \lambda_1(\boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{e}_1).$$

Thus, if $(\lambda_1, \boldsymbol{e}_1)$ is an eigenvalue-eigenvector pair for $\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1/2}$, then $(\lambda_1, \boldsymbol{f}_1)$ with $\boldsymbol{f}_1$ the normalized form of $\boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{e}_1$ is an eigenvalue-eigenvector pair for $\boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1/2}$. We have shown that $U_1 = \boldsymbol{e}_1'\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{X}_1$ and $V_1 = \boldsymbol{f}_1'\boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{X}_2$ are the first pair of canonical variables and that their correlation is $\rho_1^* = \sqrt{\lambda_1}$. Also, $\text{Var}(U_1) = \boldsymbol{e}_1'\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{\Sigma}_{11}\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{e}_1 = \boldsymbol{e}_1'\boldsymbol{e}_1 = 1$, and similarly, $\text{Var}(V_1) = 1$.
For other canonical variables, note that U_1 and an arbitrary linear combination $\boldsymbol{a}'\boldsymbol{X}_1$ are uncorrelated if

$$\text{Cov}(U_1, \boldsymbol{c}'\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{X}_1) = \boldsymbol{e}_1'\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{\Sigma}_{11}\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{c} = \boldsymbol{e}_1'\boldsymbol{c} = 0.$$

This says that $\boldsymbol{c}$ must be a linear combination of other eigenvectors. The Result is then proved by induction.

In practice, the sample covariance matrices are used to preform canonical correlation analysis. For demonstration, we use the command **cancor** in **R**.

```
Use data in Table 9-12 for demonstration:
X: 3-dimensional (3 indices of sales)
Y: 4-dimensional (4 test scores)

> setwd("C:/teaching/ama")
> da=read.table("T9-12.DAT")
> x=da[,1:3]
> y=da[,4:7]
> m1=cancor(x,y)
> m1
$cor
[1] 0.9944827 0.8781065 0.3836057

$xcoef
          [,1]        [,2]        [,3]
V1 0.008911125  0.02486719 -0.05387899
V2 0.002989377 -0.03459487  0.01478786
V3 0.011179739  0.03404200  0.05477358
```

```
$ycoef
          [,1]         [,2]         [,3]         [,4]
V4 0.009964020  0.027484475  0.035222369 -0.002672241
V5 0.004391186 -0.028796340 -0.020270754 -0.047626303
V6 0.012794882  0.070823323 -0.040032008 -0.007549523
V7 0.008975711 -0.009759438  0.001618942  0.013413863

$xcenter
     V1      V2      V3
 98.836 106.622 102.810

$ycenter
   V4    V5    V6    V7
11.22 14.18 10.56 29.76

> names(m1)
[1] "cor"     "xcoef"   "ycoef"   "xcenter" "ycenter"
> xe=as.matrix(m1$xcoef)
> ye=as.matrix(m1$ycoef)
> x1=as.matrix(x)%*%xe
> y1=as.matrix(y)%*%ye

> cor(x1[,1],y1[,1])
[1] 0.9944827
> cor(x1[,1],y1[,2])
[1] -4.615906e-17
> cor(x1[,2],y1[,2])
[1] 0.8781065

**** Monthly return example ****
> da=read.table("m-pca5c-9003.txt",header=T)
> dim(da)
[1] 168   6
> da[1,]
   IBM    HPQ   INTC    MER    MWD     Date
1 4.67 -5.716 13.534 -14.31 -8.053 19900132
> rtn=da[,1:5]
>
> y=rtn[3:168,]
> dim(y)
[1] 166   5
```

```
> x=cbind(rtn[2:167,],rtn[1:166,])
> m1=cancor(y,x)
> m1
$cor
[1] 0.39119136 0.30481247 0.25402708 0.18154150 0.07834179

$xcoef
              [,1]         [,2]         [,3]          [,4]         [,5]
IBM   0.0053887263 -0.004269681 -0.001000015 -0.0065101195  0.002392706
HPQ  -0.0043400231  0.002609529 -0.005990188 -0.0003261752  0.003725509
INTC  0.0004529081 -0.003974778  0.002373085  0.0057694043  0.001015174
MER  -0.0083384409 -0.005347687  0.005309711 -0.0059375925 -0.002147012
MWD   0.0066755480  0.008872371  0.002297098  0.0034440229  0.003386458

$ycoef
              [,1]          [,2]         [,3]          [,4]          [,5]
IBM   7.093696e-05  0.0054261528 -0.004869641 -0.0035914586 -0.0008015792
HPQ   2.091368e-03 -0.0022923598  0.002094393 -0.0009785679 -0.0026271956
INTC -1.609911e-03  0.0020331625  0.001005228  0.0013882126 -0.0010123765
MER   4.013613e-03  0.0039849073 -0.003627357  0.0065314089 -0.0010386452
MWD  -6.553790e-03 -0.0051530995  0.004320950 -0.0021044486 -0.0013036674
IBM   1.086309e-03 -0.0037942448 -0.003748766  0.0018475224  0.0016484047
HPQ  -1.180479e-03 -0.0007379369  0.001645184  0.0037698497 -0.0003217778
INTC -3.703776e-03  0.0027015932  0.001065161 -0.0035042915  0.0020564566
MER   2.193977e-03  0.0056674518  0.001413166  0.0066090999  0.0046649669
MWD  -3.189003e-03 -0.0045927382 -0.006052440 -0.0052326360 -0.0052364314
              [,6]          [,7]          [,8]          [,9]         [,10]
IBM   3.527797e-05  0.0014232820 -3.066277e-03 -0.0012330481  4.468027e-03
HPQ  -4.383624e-03  0.0048467732  2.571379e-03  0.0023527814  5.521864e-04
INTC -2.041720e-03 -0.0046596183 -1.730391e-03 -0.0003902201 -4.294113e-03
MER   1.399838e-03 -0.0004626628  6.880089e-03 -0.0061702680  3.005950e-04
MWD   2.925065e-03 -0.0015308984 -4.441831e-03  0.0043552411  3.797530e-03
IBM  -5.408307e-03 -0.0048721045 -8.487842e-04 -0.0019878989  2.596935e-03
HPQ  -1.588266e-04  0.0054342420 -4.069043e-03 -0.0030748904 -1.462014e-03
INTC -8.088000e-04 -0.0001839902  4.085529e-03 -0.0006524503  5.259674e-05
MER  -1.671939e-03 -0.0015030390  5.181337e-04  0.0064182621  4.330848e-03
MWD   3.642271e-03  0.0016987600  9.775464e-05  0.0007888860 -4.802125e-03

$xcenter
      IBM       HPQ      INTC       MER       MWD
0.9163976 1.1028133 1.9683133 1.9820663 1.6534036
```

```
$ycenter
     IBM1      HPQ1      INTC1      MER1      MWD1
0.9405422 1.0469699 2.0070181 1.9622289 1.6433554
     IBM2      HPQ2      INTC2      MER2      MWD2
0.9605181 1.0281265 2.0775000 1.8997892 1.5903614

> ye=m1$ycoef
> dim(ye)
[1] 10 10
> xe=m1$xcoef
> dim(xe)
[1] 5 5
> port=as.matrix(y)%*%xe
> xvar=as.matrix(x)%*%ye
> plot(xvar[,1],port[,1])

> v1=t(xe)%*%xe   % Normalize the eigenvector.
> v1
              [,1]          [,2]          [,3]          [,4]          [,5]
[1,]  1.621718e-04  6.768554e-05 -7.256785e-06  4.144837e-05  3.769382e-05
[2,]  6.768554e-05  1.481554e-04 -2.880827e-05  6.632190e-05  3.699810e-05
[3,] -7.256785e-06 -2.880827e-05  7.598360e-05 -1.460278e-06 -2.592113e-05
[4,]  4.144837e-05  6.632190e-05 -1.460278e-06  1.228904e-04  1.347610e-05
[5,]  3.769382e-05  3.699810e-05 -2.592113e-05  1.347610e-05  3.671279e-05
> z=port[,1]/sqrt(v1[1,1])
> mean(z)
[1] -0.3491666
> var(z)
[1] 37.37151
```

2 Large Sample Inferences

An important application of canonical correlation analysis in statistics is to test the null hypothesis $H_o : \mathbf{\Sigma}_{12} = \mathbf{0}$, i.e. the random vectors $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are indepndent under the normality assumption. The theoretical foundation for such a test is as follows:

Result 10.3. Let $\boldsymbol{X}_i = (\boldsymbol{X}'_{i,1}, \boldsymbol{X}'_{i,2})'$ be a random vector following the distribution $N_{p+q}(\boldsymbol{\mu}, \mathbf{\Sigma})$, where the dimensions of $\boldsymbol{X}_{i,1}$ and $\boldsymbol{X}_{i,2}$ are p and q, respectively, and

$$\mathbf{\Sigma} = \left[\begin{array}{cc} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{array} \right]$$

with $\mathbf{\Sigma}_{jj}$ being the covariance matrix of $\boldsymbol{X}_{i,j}$. Assume that $\mathbf{\Sigma}$ is positive definite and $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ are random sample from the normal popuation. Then the likelihood ratio test of $H_o : \mathbf{\Sigma}_{12} = \mathbf{0}$ versus $H_a : \mathbf{\Sigma}_{12} \neq \mathbf{0}$ rejects the null for large values of

$$-2\ln(\Lambda) = n \ln\left(\frac{|\boldsymbol{S}_{11}| \times |\boldsymbol{S}_{22}|}{|\boldsymbol{S}|}\right) = -n \sum_{i=1}^{p} \ln[1 - (\hat{\rho}_i^*)^2],$$

where

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} \\ \boldsymbol{S}_{21} & \boldsymbol{S}_{22} \end{bmatrix}$$

is maximum likelihood estimate of $\mathbf{\Sigma}$. In addition, for large n, the test statistic $-2\ln(\Lambda)$ is approximately distributed as a chi-square random variable with pq degrees of freedom.

Proof: Under H_o, the MLE of $\mathbf{\Sigma}$ is

$$\boldsymbol{S}_o = \begin{bmatrix} \boldsymbol{S}_{11} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{S}_{22} \end{bmatrix}.$$

Under H_a, the MLE of $\mathbf{\Sigma}$ is $\boldsymbol{S}$. Consequently,

$$-2\ln(\Lambda) = n \ln\left(\frac{|\boldsymbol{S}_{11}| \times |\boldsymbol{S}_{22}|}{|\boldsymbol{S}|}\right).$$

(Note that $\boldsymbol{S}$ can be replaced by the sample covriance matrix because of the ratio involved.) Next, using property of matrix

$$|\boldsymbol{S}| = |\boldsymbol{S}_{22}| \times |\boldsymbol{S}_{11} - \boldsymbol{S}_{12}\boldsymbol{S}_{22}^{-1}\boldsymbol{S}_{21}| = |\boldsymbol{S}_{22}| \times |\boldsymbol{S}_{11}| \times |\boldsymbol{I} - \boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12}\boldsymbol{S}_{22}^{-1}\boldsymbol{S}_{21}|.$$

Thus,

$$-2\ln(\Lambda) = n \ln\left(\frac{1}{|\boldsymbol{I} - \boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12}\boldsymbol{S}_{22}^{-1}\boldsymbol{S}_{21}|}\right) = -n \ln(|\boldsymbol{I} - \boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12}\boldsymbol{S}_{22}^{-1}\boldsymbol{S}_{21}|).$$

Next, we make use of the following three properties of determinant. (a) If λ is an eigenvalue of the matrix $\boldsymbol{A}$, then $1-\lambda$ is an eigenvalue of $\boldsymbol{I} - \boldsymbol{A}$. (b) $|\boldsymbol{A}| = \prod_{i=1}^{p} \lambda_i$, where λ_i are the eigenvalues of the matirx $\boldsymbol{A}$. Also, (c)

$$|\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12}\boldsymbol{S}_{22}^{-1}\boldsymbol{S}_{21}| = |\boldsymbol{S}_{11}^{-1/2}\boldsymbol{S}_{12}\boldsymbol{S}_{22}^{-1}\boldsymbol{S}_{21}\mathbf{\Sigma}_{11}^{-1/2}|.$$

Consequently,

$$-2\ln(\Lambda) = -n \ln\left(\prod_{i=1}^{p}[1 - (\hat{\rho}_i^*)^2]\right) = -n \sum_{i=1}^{p} \ln[1 - (\hat{\rho}_i^*)^2].$$

For finite sample, Bartlett suggests the modification

$$-\left(n - 1 - \frac{1}{2}(p + q + 1)\right) \sum_{i=1}^{p} \ln[1 - (\hat{\rho}_i^*)^2],$$

as the test statistic for H_o.

If H_o is rejected, it is natural to examine the significance of individual canonical correlations. Since canonical correlations are orderred, this leads to a sequence of hypotheses as

$$H_o^k : \rho_k^* \neq 0, \rho_{k+1}^* = \cdots = \rho_p^* = 0 \quad \text{vs} \quad H_a^k : \rho_{k+1}^* \neq 0.$$

The test statistic is

$$-\left(n - 1 - \frac{1}{2}(p+q+1)\right) \sum_{i=k+1}^{p} \ln[1 - (\hat{\rho}_i^*)^2],$$

which follows asymptotically a chi-square distribution with $(p-k)(q-k)$ degrees of freedom.

The above two test statistics are used by Johansen (1988) to test for co-integration is multivariate time series analysis.