**MODERN COLLEGE OF ARTS,SCI. & COMM. PUNE-05.**

**DEPARTMENT OF STATISTICS.**

**M.Sc.( I )**

**ST- 28**

**Practical No. 15)**

**Title : Cannonical Correlation.**

**Wadekar Amol Balasaheb (2002775)**

Q.1 Consider covariance matrix.



Verify that 1st pair of cannonical variates are U1 = 

has cannonical correlation 0.95.

Q.2 Consider population. correlation. matrix.

1 0.4 0.5 0.6

ρ = 0.4 1 0.3 0.4

0.5 0.3 1 0.2

0.6 0.4 0.2 1

Find the 1st pair of cannonical variate and canonical correlation

Q.3 The 2x1 random vectors Y and X have the joint mean vector and joint covariance matrix as

 = ,  = 



a) Calculate the cannonical corr. and .

b) Determine the cannonical variate pairs (U1, V1) and (U2, V2).

c) Let U= [(U1 ,U2)’ and (V1, V2)’.

From first principles, evaluate E 

Compare your results with the properties of cannonical variates.

Q.4 A random sample of n = 70 families were served to determine the association between certain demographic variables and certain consumption variables. Let the criterion set Y1 : Annual frequency of dinning at a restaurant and Y2 : Annual frequency of attending movies.

Predictor set is,

X1 : Age of head of household

X2 : Annual family income.

X3 : Education level of head of household.

The 70 observations on the preceding variables gave the following sample correlation matrix.

= 

a) Find the sample cannonical correlation.

b) Determine the first cannonical pair (U1, V1) and interprete these quantities.

THEORY

Canonical correlation analysis is used to identify and measure the association among two sets of variables.

Canonical correlation analysis focuses on the correlation between a linear combination of the variables in one set and a linear combination of variables in another set.

Similar to multiple regression, But with multiple dependent variables .

All other multivariate technique can be derived from this.

Let 1st group of p-variates is represented by () random vector X(1)  and 2nd group of q-variates is represented by () random vector X(2)

Mean vector =

Var-Cov matrix is,

Canonical variate pair is,

U1 = a1’X(1)

a’1 = e1’

V1 = b1’X(2)

b’1 = f1’

Var(U1) = a’\* Σ11\* a

Var(V1) = b’\* Σ22\* b

Cov(U,V) = a’\* Σ12\* b

Corr(U,V) =

ALGORITHM

1. Let 1st canonical pair is (U1,V1) and 2nd canonical pair is (U2,V2)

U1 = a1’X(1)  , U2 = a2’X(1)

V1 = b1’X(2)  , V2 = b2’X(2)

1. Find eigen vector and eigen values for A,

Let , 𝜆1 is largest eigen value and e1 is eigen vector of 𝜆1

؞ Find, a1’ = e1’

؞ U1 = a1’X(1) = X1(1) + X2(1)

1. Find eigen vector and eigen values for A,

Let , 𝜆1 is largest eigen value and f1 is eigen vector of 𝜆1

؞ Find, b1’ = f1’

؞ V1 = b1’X(2) = X1(2) + X2(2)

1. Canonical correlation = → largest eigen value.
2. Find Var(U) , Var(V) , Cov(U, V) , Corr(U, V)

**1.**

>> S = [100 0 0 0;0 1 0.95 0;0 0.95 1 0;0 0 0 100]

S =

100.0000 0 0 0

0 1.0000 0.9500 0

0 0.9500 1.0000 0

0 0 0 100.0000

>> S11 = S(1:2,1:2)

S11 =

100 0

0 1

>> S22 = S(3:4,3:4)

S22 =

1 0

0 100

>> S12 = S(1:2,3:4)

S12 =

0 0

0.9500 0

>> S21 = S(3:4,1:2)

S21 =

0 0.9500

0 0

>> A = (inv(S11)^(1/2))\*(S12)\*(inv(S22))\*(S21)\*(inv(S11)^(1/2))

A =

0 0

0 0.9025

>> [v1 d1] = eig(A)

v1 =

1 0

0 1

d1 =

0 0

0 0.9025

>> %Thus the largest eigen value of matrix A = 0.9025

>> %Therefore lambda1 = 0.9025 and eigen vector of lambda1 is e1

>> e1 =v1(:,2)

e1 =

0

1

>> a1 = (e1') \* (inv(S11)^(1/2))

a1 =

0 1

>> %Therefore U1 = X2(1)

>> B = (inv(S22)^(1/2))\*(S21)\*(inv(S11))\*(S12)\*(inv(S22)^(1/2))

B =

0.9025 0

0 0

>> [v2 d2] = eig(B)

v2 =

0 1

1 0

d2 =

0 0

0 0.9025

>> %Thus the largest eigen value of matrix B = 0.9025

>> %Therefore lambda1 = 0.9025 and eigen vector of lambda1 is f1

>> f1 =v2(:,2)

f1 =

1

0

>> b1 = (f1') \* (inv(S22)^(1/2))

b1 =

1 0

>> %Therefore V1 = X1(2)

>> %Largest eigen value = 0.95

>> c = sqrt(0.9025)

c =

0.9500

>> %Therefore canonical correlation between u1 and v1 is 0.95

%Proved

**2.**

>> %2

>> R = [1 0.4 0.5 0.6;0.4 1 0.3 0.4;0.5 0.3 1 0.2;0.6 0.4 0.2 1]

R =

1.0000 0.4000 0.5000 0.6000

0.4000 1.0000 0.3000 0.4000

0.5000 0.3000 1.0000 0.2000

0.6000 0.4000 0.2000 1.0000

>> R11 = R(1:2,1:2)

R11 =

1.0000 0.4000

0.4000 1.0000

>> R22 = R(3:4,3:4)

R22 =

1.0000 0.2000

0.2000 1.0000

>> R12 = R(1:2,3:4)

R12 =

0.5000 0.6000

0.3000 0.4000

>> R21 = R(3:4,1:2)

R21 =

0.5000 0.3000

0.6000 0.4000

>> A = (inv(R11)^(1/2))\*(R12)\*(inv(R22))\*(R21)\*(inv(R11)^(1/2))

A =

0.4370 0.2178

0.2178 0.1097

>> [v1 d1] = eig(A)

v1 =

0.8947 -0.4468

0.4468 0.8947

d1 =

0.5457 0

0 0.0009

>> %Thus the largest eigen value of matrix A = 0.5457

>> %Therefore lambda1 = 0.5457 and eigen vector of lambda1 is e1

>> e1 =v1(:,1)

e1 =

0.8947

0.4468

>> a1 = (e1') \* (inv(R11)^(1/2))

a1 =

0.8560 0.2777

>> %Therefore U1 =0.8560\*Z1(1)+0.2227\*Z2(1)

>> B = (inv(R22)^(1/2))\*(R21)\*(inv(R11))\*(R12)\*(inv(R22)^(1/2))

B =

0.2077 0.2644

0.2644 0.3389

>> [v2 d2] = eig(B)

v2 =

-0.7877 0.6161

0.6161 0.7877

d2 =

0.0009 0

0 0.5457

>> %Thus the largest eigen value of matrix B = 0.5457

>> %Therefore lambda1 = 0.5457 and eigen vector of lambda1 is f1

>> f1 =v2(:,2)

f1 =

0.6161

0.7877

>> b1 = (f1') \* (inv(R22)^(1/2))

b1 =

0.5448 0.7366

>> %Therefore V1 = 0.5448\*Z1(2)+0.7366\*Z2(2)

>> %Largest eigen value is 0.5457

>> c = sqrt(0.5457)

c =

0.7387

>> %Therefore canonical correlation between u1 and v1 is 0.7387

**3.**

>> mu = [-3;2;0;1]

mu =

-3

2

0

1

>> S = [8 2 3 1;2 5 -1 3;3 -1 6 -2;1 3 -2 7]

S =

8 2 3 1

2 5 -1 3

3 -1 6 -2

1 3 -2 7

>> S11 = S(1:2,1:2)

S11 =

8 2

2 5

>> S22 = S(3:4,3:4)

S22 =

6 -2

-2 7

>> S12 = S(1:2,3:4)

S12 =

3 1

-1 3

>> S21 = S(3:4,1:2)

S21 =

3 -1

1 3

>> %1

>> A = (inv(S11)^(1/2))\*(S12)\*(inv(S22))\*(S21)\*(inv(S11)^(1/2))

A =

0.2756 -0.0322

-0.0322 0.2690

>> [v1 d1] = eig(A)

v1 =

0.7422 0.6702

-0.6702 0.7422

d1 =

0.3046 0

0 0.2400

>> %Thus the largest eigen value of matrix A = 0.3046

>> %Therefore lambda1 = 0.3046 and eigen vector of lambda1 is e1

>> e1 =v1(:,1)

e1 =

0.7422

-0.6702

>> a1 = (e1') \* (inv(S11)^(1/2))

a1 =

0.3168 -0.3622

>> %Therefore U1 =0.3168\*x1(1)-0.3622\*X2(1)

>> B = (inv(S22)^(1/2))\*(S21)\*(inv(S11))\*(S12)\*(inv(S22)^(1/2))

B =

0.2946 -0.0234

-0.0234 0.2500

>> [v2 d2] = eig(B)

v2 =

0.9193 0.3936

-0.3936 0.9193

d2 =

0.3046 0

0 0.2400

>> %Thus the largest eigen value of matrix B = 0.3046

>> %Therefore lambda1 = 0.3046 and eigen vector of lambda1 is f1

>> f1 =v2(:,1)

f1 =

0.9193

-0.3936

>> b1 = (f1') \* (inv(S22)^(1/2))

b1 =

0.3647 -0.0951

>> %Therefore V1 = 0.3647\*X1(2)-0.0951\*X2(2)

>>

>> %largest eigen value 0.3046

>> c = sqrt(0.3046)

c =

0.5519

>> %Therefore canonical correlation between u1 and v1 is 0.5519

>> %Similarly calculate U2 and V2 for 2nd eigen value

>> %2nd eigen value of A = 0.2400

>> %Therefore lambda2 = 0.2400 and eigen vector of lambda2 is e2

>> e2 =v1(:,2)

e2 =

0.6702

0.7422

>> a2 = (e2') \* (inv(S11)^(1/2))

a2 =

0.1962 0.3017

>> %Therefore U2 =0.1962\*x1(1)+0.3017\*X2(1)

>> %2nd eigen value of B = 0.2400

>> %Therefore lambda2 = 0.2400 and eigen vector of lambda2 is f2

>> f2 =v2(:,2)

f2 =

0.3936

0.9193

>> b2 = (f2') \* (inv(S22)^(1/2))

b2 =

0.2263 0.3858

>> %Therefore V2 = 0.2263\*X1(2)+0.3858\*X2(2)

>> c = sqrt(0.2400)

c =

0.4899

>> %Therefore canonical correlation between u1 and v1 is 0.4899

ii)

>> A = (inv(S11)^(1/2))\*(S12)\*(inv(S22))\*(S21)\*(inv(S11)^(1/2))

A =

0.2756 -0.0322

-0.0322 0.2690

>> [v1 d1] = eig(A)

v1 =

0.7422 0.6702

-0.6702 0.7422

d1 =

0.3046 0

0 0.2400

>> %Thus the largest eigen value of matrix A = 0.3046

>> %Therefore lambda1 = 0.3046 and eigen vector of lambda1 is e1

>> e1 =v1(:,1)

e1 =

0.7422

-0.6702

>> a1 = (e1') \* (inv(S11)^(1/2))

a1 =

0.3168 -0.3622

>> %Therefore U1 =0.3168\*x1(1)-0.3622\*X2(1)

>> B = (inv(S22)^(1/2))\*(S21)\*(inv(S11))\*(S12)\*(inv(S22)^(1/2))

B =

0.2946 -0.0234

-0.0234 0.2500

>> [v2 d2] = eig(B)

v2 =

0.9193 0.3936

-0.3936 0.9193

d2 =

0.3046 0

0 0.2400

>> %Thus the largest eigen value of matrix B = 0.3046

>> %Therefore lambda1 = 0.3046 and eigen vector of lambda1 is f1

>> f1 =v2(:,1)

f1 =

0.9193

-0.3936

>> b1 = (f1') \* (inv(S22)^(1/2))

b1 =

0.3647 -0.0951

>> %Therefore V1 = 0.3647\*X1(2)-0.0951\*X2(2)

>> %Similarly calculate U2 and V2 for 2nd eigen value

>> %2nd eigen value of A = 0.2400

>> %Therefore lambda2 = 0.2400 and eigen vector of lambda2 is e2

>> e2 =v1(:,2)

e2 =

0.6702

0.7422

>> a2 = (e2') \* (inv(S11)^(1/2))

a2 =

0.1962 0.3017

>> %Therefore U2 =0.1962\*x1(1)+0.3017\*X2(1)

>> %2nd eigen value of B = 0.2400

>> %Therefore lambda2 = 0.2400 and eigen vector of lambda2 is f2

>> f2 =v2(:,2)

f2 =

0.3936

0.9193

>> b2 = (f2') \* (inv(S22)^(1/2))

b2 =

0.2263 0.3858

>> %Therefore V2 = 0.2263\*X1(2)+0.3858\*X2(2)

**iii)**

>> %3

>> %To find mean vector of U=[U1 U2] and V=[V1 V2]

>> %E(U) = a'mu1 E(V) = b'mu2

>> a = [a1;a2]

a =

0.3168 -0.3622

0.1962 0.3017

>> b = [b1;b2]

b =

0.3647 -0.0951

0.2263 0.3858

>> mu1 = mu(1:2)

mu1 =

-3

2

>> mu2 = mu(3:4)

mu2 =

0

1

>> m1 = a\*mu1

m1 =

-1.6749

0.0146

>> m2 = b\*mu2

m2 =

-0.0951

0.3858

>> m = [m1;m2]

m =

-1.6749

0.0146

-0.0951

0.3858

>> %m is vector of mean of U and V

>> %Now we calculate cov(U,V)

>> %To calculate cov(U,V) we find Suu,Suv,Svu,Svv

>> Suu = a\*S11\*a'

Suu =

1.0000 -0.0000

-0.0000 1.0000

>> Suv = a\*S12\*b'

Suv =

0.5519 0.0000

0.0000 0.4899

>> Svu = b\*S21\*a'

Svu =

0.5519 0.0000

0.0000 0.4899

>> Svv = b\*S22\*b'

Svv =

1.0000 0.0000

0.0000 1.0000

>> COV = [Suu Suv;Svu Svv]

COV =

1.0000 -0.0000 0.5519 0.0000

-0.0000 1.0000 0.0000 0.4899

0.5519 0.0000 1.0000 0.0000

0.0000 0.4899 0.0000 1.0000

>> %COV is the covariance matrix of U and V

**4.**

>> %4

>> R = [1 0.8 0.26 0.67 0.34;0.8 1 0.33 0.59 0.34;0.26 0.33 1 0.37 0.21;0.67 0.59 0.37 1 0.35;0.34 0.34 0.21 0.35 1]

R =

1.0000 0.8000 0.2600 0.6700 0.3400

0.8000 1.0000 0.3300 0.5900 0.3400

0.2600 0.3300 1.0000 0.3700 0.2100

0.6700 0.5900 0.3700 1.0000 0.3500

0.3400 0.3400 0.2100 0.3500 1.0000

>> R11 = R(1:2,1:2)

R11 =

1.0000 0.8000

0.8000 1.0000

>> R22 = R(3:5,3:5)

R22 =

1.0000 0.3700 0.2100

0.3700 1.0000 0.3500

0.2100 0.3500 1.0000

>> R12 = R(1:2,3:5)

R12 =

0.2600 0.6700 0.3400

0.3300 0.5900 0.3400

>> R21 = R(3:5,1:2)

R21 =

0.2600 0.3300

0.6700 0.5900

0.3400 0.3400

>> A = (inv(R11)^(1/2))\*(R12)\*(inv(R22))\*(R21)\*(inv(R11)^(1/2))

A =

0.3221 0.2084

0.2084 0.1861

>> [v1 d1] = eig(A)

v1 =

0.8094 -0.5872

0.5872 0.8094

d1 =

0.4733 0

0 0.0349

>> %Thus the largest eigen value of matrix A = 0.4733

>> %Therefore lambda1 = 0.4733 and eigen vector of lambda1 is e1

>> e1 =v1(:,1)

e1 =

0.8094

0.5872

>> a1 = (e1') \* (inv(R11)^(1/2))

a1 =

0.7689 0.2721

>> %Therefore U1 =0.7689\*Z1(1)+0.2721\*Z2(1)

>> B = (inv(R22)^(1/2))\*(R21)\*(inv(R11))\*(R12)\*(inv(R22)^(1/2))

B =

0.0531 0.0881 0.0461

0.0881 0.3962 0.1451

0.0461 0.1451 0.0589

>> [v2 d2] = eig(B)

v2 =

-0.2288 -0.9001 -0.3708

-0.9105 0.3326 -0.2456

-0.3444 -0.2814 0.8956

d2 =

0.4733 0 0

0 0.0349 0

0 0 0.0000

>> %Thus the largest eigen value of matrix B = 0.4733

>> %Therefore lambda1 = 0.4733 and eigen vector of lambda1 is f1

>> f1 =v2(:,1)

f1 =

-0.2288

-0.9105

-0.3444

>> b1 = (f1') \* (inv(R22)^(1/2))

b1 =

-0.0491 -0.8975 -0.1900

>> %Therefore V1 = -0.0491\*Z1(2)-0.8975\*Z2(2)-0.1900\*Z3(2)

>> %Largest eigen value is 0.4733

>> c = sqrt(0.4733)

c =

0.6880

>> %Therefore sample canonical correlation between u1 and v1 is 0.6880

**ii)**

>> A = (inv(R11)^(1/2))\*(R12)\*(inv(R22))\*(R21)\*(inv(R11)^(1/2))

A =

0.3221 0.2084

0.2084 0.1861

>> [v1 d1] = eig(A)

v1 =

0.8094 -0.5872

0.5872 0.8094

d1 =

0.4733 0

0 0.0349

>> %Thus the largest eigen value of matrix A = 0.4733

>> %Therefore lambda1 = 0.4733 and eigen vector of lambda1 is e1

>> e1 =v1(:,1)

e1 =

0.8094

0.5872

>> a1 = (e1') \* (inv(R11)^(1/2))

a1 =

0.7689 0.2721

>> %Therefore U1 =0.7689\*Z1(1)+0.2721\*Z2(1)

>> B = (inv(R22)^(1/2))\*(R21)\*(inv(R11))\*(R12)\*(inv(R22)^(1/2))

B =

0.0531 0.0881 0.0461

0.0881 0.3962 0.1451

0.0461 0.1451 0.0589

>> [v2 d2] = eig(B)

v2 =

-0.2288 -0.9001 -0.3708

-0.9105 0.3326 -0.2456

-0.3444 -0.2814 0.8956

d2 =

0.4733 0 0

0 0.0349 0

0 0 0.0000

>> %Thus the largest eigen value of matrix B = 0.4733

>> %Therefore lambda1 = 0.4733 and eigen vector of lambda1 is f1

>> f1 =v2(:,1)

f1 =

-0.2288

-0.9105

-0.3444

>> b1 = (f1') \* (inv(R22)^(1/2))

b1 =

-0.0491 -0.8975 -0.1900

>> %Therefore V1 = -0.0491\*Z1(2)-0.8975\*Z2(2)-0.1900\*Z3(2)