Multivariate Analysis
Assignment 7

1) By definition, squared canonical correlation is the max. squared correlation between linear combinations of y's & x's.

$U = a'y$ & $v = b'x$ (let)

Task : Find $r^2 = \max_{a,b} r^2_{u,v} = \max_{a,b} \frac{(a^T S_{yx} b)^2}{(a^T S_{yy} a)(b^T S_{xx} b)}$

Now, $\frac{\partial}{\partial a} r^2_{u,v} = 0$

$$\Rightarrow \frac{(a^T S_{yy} a)(b^T S_{xx} b) . 2(a^T S_{yx} b)(S_{yx} b) - (a^T S_{yx} b)^2 2(S_{yy} a)(b^T S_{xx} b)}{\{(a^T S_{yy} a)(b^T S_{xx} b)\}^2}$$

$= 0$

$$\Rightarrow \frac{(a^T S_{yy} a)(S_{yx} b) - (a^T S_{yx} b)(S_{yy} a)}{(a^T S_{yy} a)(b^T S_{xx} b)} = 0$$

$$\Rightarrow (S_{yx} b) - \left(\frac{a^T S_{yx} b}{a^T S_{yy} a}\right)(S_{yy} a) = 0 \quad \text{— (i)}$$

And $\frac{\partial}{\partial b} r^2_{u,v} = 0$

$$\Rightarrow \frac{(a^T S_{yy} a)(b^T S_{xx} b) . 2(a^T S_{yx} b)(S_{xy} a) - (a^T S_{yx} b)^2 (a^T S_{yy} a) 2 . (S_{xx} b)}{\{(a^T S_{yy} a)^2 (b^T S_{xx} b)\}^2}$$

$= 0$

$$\Rightarrow (S_{xy} a) - \left(\frac{a^T S_{yx} b}{b^T S_{xx} b}\right)(S_{xx} b) = 0 \quad \text{— (ii)}$$

from (i) & (ii) $\rightarrow$ $a = \frac{S_{yy}^{-1} S_{yx} b}{rp}$ where $p = \sqrt{\frac{b^T S_{xx} b}{a^T S_{yy} a}}$

$$\Rightarrow \frac{S_{xy} S_{yy}^{-1} S_{yx} b}{rp} - \frac{r}{p}(S_{xx} b) = 0$$

$$\Rightarrow S_{xy} S_{yy}^{-1} S_{yx} b - r^2 S_{xx} b = 0 \Rightarrow (S_{xx}^{-1} S_{xy} S_{yy}^{-1} S_{yx} - r^2 I) b = 0 \quad \text{(iii)}$$

Similarly: from $b = (S_{xx}^{-1} S_{xy} a) P/r$, we get

$$(S_{yx} S_{xx}^{-1} S_{xy} - r^2 S_{yy}) a = 0$$

$\Rightarrow \left(S_{yy}^{-1} S_{yx} S_{xx}^{-1} S_{xy} - r^2 I\right) a = 0$ — (iv)

- Canonical correlation & coefficient vectors are the eigenvalues & eigenvectors of eigenvalue problems in (iii) & (iv).

let A, B be two matrices s.t $AB v = \lambda v$

$\Rightarrow BA(Bv) = \lambda(Bv)$ $\Rightarrow$ AB & BA have the same eigenvalue λ (with different corresponding eigenvectors).

Assign: $A := S_{yy}^{-1} S_{yx}$, $B := S_{xx}^{-1} S_{xy}$ $\xrightarrow{\text{from (iii) \& (iv)}}$

$(BA - r^2 I)\, b = 0$ & $(AB - r^2 I) a = 0$

which solve to yield the same value of r^2 (eigenvalue of AB, BA principle).

2) Using the relations:

- $a = (S_{yy}^{-1} S_{yx}\, b)\, 1/rp$ where $p = \sqrt{\dfrac{b^T S_{xx} b}{a^T S_{yy}\, a}}$
- $b = (S_{xx}^{-1} S_{xy}\, a)\, p/r$

$S_u^2 = a^T S_{yy}\, a = 1$ & $S_v^2 = b^T S_{xx}\, b = 1$ [Given]

$\therefore p = 1$

$\rightarrow a = \dfrac{1}{r}(S_{yy}^{-1} S_{yx}\, b)$

$\rightarrow b = \dfrac{1}{r}(S_{xx}^{-1} S_{xy}\, a)$ {Proved}

3) Eigenvalue problem: $(CC^T - \lambda_i I)p_i = 0$

$- \lambda_i$: eigenvalue, p_i : eigenvector.

Using the given value of C from the problem.

$$(S_{yy}^{-1/2} S_{yx} S_{xx}^{-1/2} S_{xx}^{-1/2} S_{xy} S_{yy}^{-1/2} - \lambda_i I) p_i = 0 \quad \text{— (i)}$$

— S_{xx}, S_{yy}, S_{xy} are symmetric.

set $a_i = S_{yy}^{-1/2} p_i$. Pre-multiplying (i) by $S_{yy}^{-1/2}$ →

$$(S_{yy}^{-1} S_{yx} S_{xx}^{-1} S_{xy} S_{yy}^{-1/2} - \lambda_i S_{yy}^{-1/2} p_i) = 0$$

$$\rightarrow (S_{yy}^{-1} S_{yx} S_{xx}^{-1} S_{xy} - \lambda_i p_i) = 0 \quad \text{— (ii)}$$

We know that the λ_i in (ii) corresponds to the i^{th} squared canonical correlation and the corresponding coefficient vector is $a_i = S_{yy}^{-1/2} p_i$.

Also, eigenvalue problem: $(C^T C - \beta_i I) q_i = 0$

— β_i : eigenvalue, q_i : eigenvector

Using the given value of C from the problem.

$$(S_{xx}^{-1/2} S_{xy} S_{yy}^{-1/2} S_{yy}^{-1/2} S_{yx} S_{xx}^{-1/2} - \beta_i I) q_i = 0 \quad \text{— (iii)}$$

set $b_i = S_{xx}^{-1/2} q_i$, Pre-multiplying (iii) with $S_{xx}^{-1/2}$ →

$$(S_{xx}^{-1} S_{xy} S_{yy}^{-1} S_{yx} S_{xx}^{-1/2} - \beta_i S_{xx}^{-1/2} I) q_i = 0$$

$$\Rightarrow (S_{xx}^{-1} S_{xy} S_{yy}^{-1} S_{yx} - \beta_i I) b_i = 0 \quad \{\text{As } b_i = S_{xx}^{-1/2} q_i\} \quad \text{— (iv)}$$

We know: b_i in (iv) corresponds to the i^{th} squared canonical correlation with corresponding coefficient vector as $b_i = S_{xx}^{-1/2} q_i$.

4) Let $\begin{cases} U = [u_1 \quad u_2 \quad \dots \quad u_s]', \\ V = [v_1 \quad v_2 \quad \dots \quad v_s] \end{cases}$

$$C = S_{yy}^{-1/2} S_{yx} S_{xx}^{-1/2} \quad \{Given\}$$

Singular Value Decomposition on C yields $C = PDQ^T$

: P is a $p \times p$ matrix — of normalized eigenvectors of CC^T

: Q is a $q \times p$ matrix — of first p normalized eigenvectors of C^TC

: D is a diagonal matrix — of $r_i(s)$ s.t r_i = square root of the i^{th} largest eigenvalue of CC^T $\forall\, i = 1(1)p$

[Assuming $p \le q$.]

Using: $u_i = a_i Y, = S_{yy}^{-1/2} P_i Y$

We have $U = AY$ with $A = P^T S_{yy}^{-1/2}$

And: $V = BX$ with $B = Q^T S_{xx}^{-1/2}$

- $A S_{yy} A^T = P^T S_{yy}^{-1/2} S_{yy} S_{yy}^{-1/2} P = P^T P = I$

  [Orthonormality of matrix of eigenvectors]
- Similarly, $B S_{xx} B^T = Q^T S_{xx}^{-1/2} S_{xx} S_{xx}^{-1/2} Q = Q^T Q = I$
- And $A S_{yx} B^T = P^T S_{yy}^{-1/2} S_{yx} S_{xx}^{-1/2} Q = P^T C Q$

$$= P^T P D Q^T Q$$
$$= (P^T P) D (Q^T Q)$$
$$= D.$$

∘ Computing $\hat{cov}(U,V) = \hat{cov}(AY, BX)$

$$= \begin{bmatrix} \hat{cov}(AY, AY) & \hat{cov}(AY, BX) \\ \hat{cov}(BX, AY) & \hat{cov}(BX, BX) \end{bmatrix} = \begin{bmatrix} A S_{yy} A^T & A S_{yx} B^T \\ B S_{xy} A^T & B S_{xx} B^T \end{bmatrix}$$

→ Combining: $\hat{cov}(U,V) = \begin{bmatrix} I & D \\ D & I \end{bmatrix}$ $\quad \left\{ \begin{array}{l} As\ (A S_{yx} B^T)^T \\ = B S_{yx} A^T \end{array} \right\}$

expanding:

$$\begin{bmatrix} \hat{cov}(u_1,u_1) & \dots & \hat{cov}(u_1,u_p) & \hat{cov}(u_1,v_1) & \dots & \hat{cov}(u_1,v_p) \\ \hat{cov}(u_2,u_1) & \dots & \hat{cov}(u_2,u_p) & \hat{cov}(u_2,v_1) & \dots & \hat{cov}(u_2,v_p) \\ \vdots & & \vdots & \vdots & & \vdots \\ cov(v_p,u_1) & \dots & cov(v_p,u_p) & \hat{cov}(v_p,v_1) & \dots & \hat{cov}(u_p,v_p) \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & \dots & 0 & r_1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & r_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & r_p \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 \\ r_1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_p & 0 & 0 & \dots & 1 \end{bmatrix}$$

Clearly,

$\hat{cov}(u_i,u_i) = \hat{cov}(v_i,v_i) = 1 \;\forall\, i = 1(1)p$

$\hat{cov}(u_i,v_j) = \hat{cov}(v_i,u_j) = 0 \;\forall\, i,j = 1(1)p$ and $i \ne j$

$\hat{cov}(u_i,v_i) = \hat{cov}(v_i,u_i) = r_i \;\forall\, i = 1(1)p$

→ All canonical variates are uncorrelated except u_i, v_i pairs $\forall\, i = 1(1)p$ {Proved}

5) Using Lagrangian scheme: $r_1^2 = \max_{a,b} (a^T S_{yx} b)^2$

s.t $a^T S_{yy} a = 1$ & $b^T S_{xx} b = 1$

We need to maximize $\underset{a,b}{Z} = \max_{a,b} (a^T S_{yx} b)^2 - \rho_1 (a^T S_{yy} a - 1) - \rho_2 (b^T S_{xx} b - 1)$

$\frac{\partial Z}{\partial a} = 0 \Rightarrow 2(a^T S_{yx} b).(S_{yx} b) - 2\rho_1 S_{yy} a = 0$

[Let $k = a^T S_{yx} b$] $\Rightarrow \cancel{2}k S_{yx} b - \cancel{2}\rho_1 S_{yy} a = 0$ —— (i)

$\frac{\partial Z}{\partial b} = 0 \Rightarrow 2(a^T S_{yx} b)(S_{xy} a) - 2\rho_2 S_{xx} b = 0$

$\Rightarrow k S_{xy} a - \rho_2 S_{xx} b = 0$ —— (ii)

$a^T \times$ (i) : $k a^T S_{yx} b - \rho_1 a^T S_{yy} a = 0 \Rightarrow \rho_1 = k^2$

$b^T \times$ (ii) : $k b^T S_{xy} a - \rho_2 b^T S_{xx} b = 0 \Rightarrow \rho_2 = k^2$

As $a^T S_{yy} a = b^T S_{xx} b = 1$ and $b^T S_{xy} a = (a^T S_{yx} b)^T = k^T = k$.

$\therefore$ $k S_{yx} b - k^2 S_{yy} a = 0$ & $k S_{xy} a - k^2 S_{xx} b = 0$ } $a = (S_{yy}^{-1} S_{yx} b) \frac{1}{k}$

and $\therefore k S_{xy} (S_{yy}^{-1} S_{yx} b) \frac{1}{k} - k^2 S_{xx} b$

$= (S_{xx}^{-1} S_{xy} S_{yy}^{-1} S_{yx} - k^2 I) = 0$ -(iii)

Similarly, we obtain: $(S_{yy}^{-1} S_{yx} S_{xx}^{-1} S_{xy} - k^2 I)a = 0$ —— (iv)

(iii) & (iv) are two eigenvalue problems.

The solution for which is known: $k^2 = \max_{a,b} (a^T S_{yx} b)^2 = r_1^2$

Thus, r_1^2 has the exact same value as we calculated in the solution to problem 1.

Problem 6

April 15, 2021

```
In [1]: # A problem on canonical correlation
```

1 *Problem* 6

```
In [2]: %%capture
        !pip install statsmodels

        # Utilizing Numpy and Statsmodels packages
        import numpy as np
        from statsmodels.multivariate.cancorr import CanCorr

In [3]: y = [[60, 69, 62],
            [56, 53, 84],
            [80, 69, 76],
            [55, 80, 90],
            [62, 75, 68],
            [74, 64, 70],
            [64, 71, 66],
            [73, 70, 64],
            [68, 67, 75],
            [69, 82, 74],
            [60, 67, 61],
            [70, 74, 78],
            [66, 74, 78],
            [83, 70, 74],
            [68, 66, 90],
            [78, 63, 75],
            [77, 68, 74],
            [66, 77, 68],
            [70, 70, 72],
            [75, 65, 71]]

        x = [[97, 69, 98],
            [103, 78, 107],
            [66, 99, 130],
            [80, 85, 114],
            [116, 130, 91],
```

```
            [109, 101, 103],
            [77, 102, 130],
            [115, 110, 109],
            [76, 85, 119],
            [72, 133, 127],
            [130, 134, 121],
            [150, 158, 100],
            [150, 131, 142],
            [99, 98, 105],
            [119, 85, 109],
            [164, 98, 138],
            [144, 71, 153],
            [77, 82, 89],
            [114, 93, 122],
            [77, 70, 109]]

        y = np.asarray(y)
        x = np.asarray(x)

        standardized_y = (y - np.mean(y, axis = 0)) / np.std(y, axis = 0)
        standardized_x = (x - np.mean(x, axis = 0)) / np.std(x, axis = 0)
```

```
In [4]: S_matrix = np.cov(y.T, x.T)
```

```
In [5]: np.set_printoptions(suppress = True)
        print('S matrix = ')
        print(np.round(S_matrix, 4))
```

```
S matrix =
[[ 61.0632  -5.0947  -5.5789  28.2368  -6.8105  37.9895]
 [ -5.0947  41.4842  -2.4737 -41.1842  66.1368  -5.4316]
 [ -5.5789  -2.4737  64.3684   9.5    -27.1053  17.8421]
 [ 28.2368 -41.1842   9.5    876.9342 268.3158 143.3684]
 [ -6.8105  66.1368 -27.1053 268.3158 621.6211  -0.0316]
 [ 37.9895  -5.4316  17.8421 143.3684  -0.0316 293.0105]]
```

```
In [6]: model = CanCorr(endog = y, exog = x)
        model_std = CanCorr(endog = standardized_y, exog = standardized_x)
```

1.1 (*a*)

```
In [7]: print('Canonical Correlations between (y_1, y_2, y_3) and (x_1, x_2, x_3): ')
        print(np.ndarray.tolist(np.round(model.cancorr, 6)))
```

```
Canonical Correlations between (y_1, y_2, y_3) and (x_1, x_2, x_3):
[0.590852, 0.309003, 0.052614]
```

```
In [8]: print('Squared Canonical Correlations between (y_1, y_2, y_3) and (x_1, x_2, x_3): ')
        print(np.ndarray.tolist(np.round(np.square(model.cancorr), 6)))
```

```
Squared Canonical Correlations between (y_1, y_2, y_3) and (x_1, x_2, x_3):
[0.349106, 0.095483, 0.002768]
```

1.2 (*b*)

```
In [9]: print('The canonical coefficients for \'endog\' i.e y:')
        print(np.round(model.y_cancoef, 6))
        print()

        print('The canonical coefficients for \'exog\' i.e x:')
        print(np.round(model.x_cancoef, 6))
```

```
The canonical coefficients for 'endog' i.e y:
[[-0.003864 -0.027247 -0.011015]
 [ 0.033479 -0.011335  0.006061]
 [-0.006893 -0.012139  0.02514 ]]

The canonical coefficients for 'exog' i.e x:
[[-0.00583   0.001246 -0.006374]
 [ 0.009126 -0.001762 -0.003533]
 [ 0.000245 -0.013754  0.002936]]
```

1.3 (*c*)

```
In [10]: print('Tests of Significance for each Canonical Correlation:')
         print()
         print(model.corr_test(), end = '')
```

```
Tests of Significance for each Canonical Correlation:

                          Cancorr results
===================================================================
  Canonical Correlation Wilks' lambda Num DF  Den DF F Value Pr > F
-------------------------------------------------------------------
0                0.5909        0.5871 9.0000 34.2229  0.9301 0.5120
1                0.3090        0.9020 4.0000 30.0000  0.3969 0.8093
2                0.0526        0.9972 1.0000 16.0000  0.0444 0.8357
-------------------------------------------------------------------

-------------------------------------------------------------------
Multivariate Statistics and F Approximations
---------------------------------------------------------------------
                         Value    Num DF    Den DF   F Value   Pr > F
---------------------------------------------------------------------
```

```
Wilks' lambda           0.5871   9.0000   34.2229    0.9301   0.5120
Pillai's trace          0.4474   9.0000   48.0000    0.9347   0.5043
Hotelling-Lawley trace  0.6447   9.0000   19.0526    0.9604   0.5000
Roy's greatest root     0.5363   3.0000   16.0000    2.8605   0.0696
===================================================================
```

All F values are lower than corrsponding critical values.
$\therefore$ *none of the correlations are significant.*

```
In [11]: # ^_^ Thank You
```