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.
∴ none of the correlations are significant.

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