Süleyman Demirel Üniversitesi
İktisadi ve İdari Bilimler
Fakültesi Dergisi
Y.2018, C.23, S.3, s.1173-1183.

Suleyman Demirel University
The Journal of Faculty of Economics
and Administrative Sciences
Y.2018, Vol.23, No.3, pp.1173-1183.

DETERMINING THE ASSOCIATION BETWEEN ECONOMIC AND SOCIAL INDICATORS BY USING CANONCAL CORRELATION ANALYSIS: THE ARABIC UNION SAMPLE

BAZI EKONOMİK VE SOSYAL GÖSTERGELER ARASINDAKİ İLİŞKİNİN KANONİK KORELASYON ANALİZİ İLE BELİRLENMESİ: ARAP BİRLİĞİ ÖRNEĞİ

Ayhan DEMİRCİ*

* Dr. Öğr. Üyesi, Toros Üniversitesi, İktisadi, İdari ve Sosyal Bilimler Fakültesi, Uluslararası Ticaret ve Lojistik Bölümü, ayhan.demirci@toros.edu.tr, https://orcid.org/0000-0003-3788-4586

ABSTRACT

Canonical correlation analysis is a multivariate statistical teqnique that can be used to determine the relationship between two multiple wariable sets. In the first part of this paper which applicated the canonical correlation method, theoretical explanation of the usage of canonical correlation analysis is exposed. Than, in the second part of the study, the applicatioan of canonical correlation analysis method is exsposed for some economic and social indicators of 20 countries that are the members of the Arabic Union. By using canonical correlation analysis method, it is shown that there is a meaningful relation between the two wariable sets. But it can be said that, there is no strong relation between two wariable sets, because of the period of choosing the wariables.

Keywords: *Canonical Correlation Analysis, Multivariable Statistics.*

Jel Codes: *C40, C46.*

ÖZ

Kanonik korelasyon analizi, çok değişkenli iki küme arasındaki ilişkiyi belirlemek için kullanılan, çok değişkenli bir istatistiksel yöntemdir. Kanonik korelasyon analizinin uygulandığı bu çalışmanın ilk bölümünde, teorik olarak yöntemin detayları ortaya konmuştur. İkinci bölümde, Arap Birliği üyesi olan 20 ülkeye ait bazı ekonomik ve sosyal göstergeler kullanılarak kanonik korelasyon analizi uygulanmıştır. Yöntemin uygulanması neticesinde, iki değişken seti arasında anlamlı bir ilişki olduğu görülmüştür. Ancak iki değişmez set arasında belirlenen kanonik korelasyon katsayısına bakıldığında, ilişkinin çok yüksek düzeyde çıkmaması, değişken seçiminden kaynalandığı söylenebilir.

Anahtar Kelimeler: *Kanonik Korelasyon Analizi, Çok Değişkenli İstatistik.*

Jel Kodları: *C40, C46.*

OVERVIEW OF CANONICAL CORRELATION ANALYSIS

Before sharing technical information about canonical correlation analysis, it is useful to know following terms to make the method understandable (Garcia-Gallego and Mures-Quintana, 2016: 252);

Standardised coefficients (canonical weights): the coefficients for the original variables in the linear combinations are their relative contribution to each canonical variate, but they can be influenced by the existence of multicollinearity, so it is more usual to interpret the structure coefficients.

Structure coefficients (canonical loadings): they measure the level in which the original variables in each set are represented by the respective canonical variates, as they are the correlation between them. The proportion of variance of the variable that is the explained by the canonical variate is the squared coefficient and the mean for each group is the total proportion of explained variance in the group. The total addition measures the proportion of variance in a set of variables that is explained by all canonical variates which are obtained for that set.

Canonical cross-loadings: they measure the correlation between the original variables in a group and the canonical variates obtained for the other one. They are computed as the product of the above structure coefficients and the corresponding canonical correlation coefficient.

Redundancy index: it was proposed as a measure of the correlation between both sets of original variables, since canonical correlations measure the correlation between the linear combinations for each set of variables, but not between the groups themselves. The squared canonical correlation –which accounts for the proportion of variance in a canonical variate of each pair that is explained by the other canonical variate in the pair– is multiplied by the variance explained by each canonical variate. Their total addition for all canonical variates is the redundancy index, which measures the proportion of variance in a set of variables that is explained by the other set.

INTRODUCTION

The simplest correlation (correlation) known in statistics is the relation called "simple correlation" between two random variables X and Y. If the number of variables is p and if the correlation between one of the variables and the remaining p-1 variables is searched, then the correlation coefficient to be calculated is called the "multiple correlation coefficient". In canonical correlation analysis, which is the most general and most complex relationship analysis, it is concerned with the relations between two variable sets drawn from a very variable population (Kalaycı, 2008:237).

Canonical correlation is a multivariate statistical model that facilities tha study of interrelationships among multiple dependent variables and multiple independent variables. The components of one set of variables are linearly related to the components of a second set of variables; it is the most general linear model with many other parametric significance as special cases (Pugh and Hu, 1991: 147).

MATHEMATICAL APPRECIATION OF CANONICAL CORRELATION ANALYSIS

Formally, canonical correlation can be written (Fornell and Larcker, 1980: 458);

$$Y_1 = a_{11} y_{11} + \dots + a_{1p} y_{1p}$$

$$X_1 = b_{11} x_{11} + \dots + b_{1q} x_{1q}$$

.

.

.

.

.

.

$$Y_i = a_{i1} y_{i1} + \dots + a_{ip} y_{ip}$$

$$X_i = b_{i1} x_{i1} + \dots + b_{iq} x_{iq}$$

where:

yjj = the criterion variables (1 < j < p)

Xjij = the predictor variables (1 < k < q)

i = tbe number of variate pairs (the smaller of p and q)

p = the number of criterion variables

q = the number of predictor variables

A Yj = canonical variate for the criterion variables in the ith pair

Xj = canonical variate for the predictor variables in the ith pair

ajj = criterion set canonical weight for the jth variable in the ith pair (1 < J < p)

bik = predictor set canonical weights for the kth variable in the ith pair (1 < k < q)

The coefficient of the simple correlation is between -1 and 1, while the canonical correlation is between 0 and 1 (Oktay and Çınar, 2002: 11-31).

Hotelling has shown that canonical analysis is the equivalent of performing independent principal components analyses on each of two sets of variables. Then the resulting component structures are rotated to develop weights for each variable that produce maximal correlations between components on each side (Alpert and Peterson, 1972: 188).

Theoretically, if a set of dependent and independent variables can be distinguished between two sets of variables, then the aim of the analysis of the canonical correlation is to determine whether the set of independent variables affects the set of dependent variables. However, in the analysis of canonical correlation, it is not compulsory to subject two sets of variables to a separate set of independent and independent variables (Kalaycı, 2008: 238).

USES OF CANONICAL CORRELATION ANALYSIS

Canonical Correlation Analysis has four goals;

a. To test whether two sets of variables that are obtained from the same set are independent of each other,

b. To detect variables in both sets that are most likely to contribute to the inter-set correlation,

c. To determine linear combinations that maximize the correlation between sets of estimates and criterion variables,

d. To estimate the other set of values by looking at the linear combination values of the individual set of variables.

By performing canonical correlation analysis via a structural relations model, two important advantages are gained. First, the problem of determining statistical significance for individual parameters in canonical correlations is overcome. For large samples, the estimated standart errors can be used to calculate critical ratios for the evaluation of individual weights and cross loadings. Second, it was shown that the structural relations model is a more flexible tool for data analysis compared to canonical correlation. It is possible that a cannonical model may fail to reject the null hypothesis regarding the relationship between two-variates, although this relationship is significant in a less restrictive structural relations model. Thus, if the data are inconsistent with the cannonical model it does not necessarily follow that one’s theory should be rejected. By altering some of the assumptions of canonical analysis, such as the requirement that tha variates be exact linear combinations of the variables, it is possible to arrive at a different conclusion (Bagozzi vd., 1981: 453-454).

There is a concern with the application of canonical correlation analysis is provided in the researc of Fornell (1979). He illustrates that canonical correlation analysis may be influenced by relatively minor portions of variance in the data. It is possible that a canonical solution may yield statistically significant results and high canonical correlation coefficients, yet lack significant explanatory power (redundancy). This is because the objective of the canonical model is to find weighted sums of the original variables that maximally correlate, while the size of the canonical loadings (and thus the redundancy) are not considered in this optimizations. Therefore, the initial correlation matrix of observed variables must be examined with respect to the distribution of high and low correlation coefficients before it is submitted to canonical correlation analysis (Fornell and Larcker, 1980: 468-469).

The nature of the canonical variates, derived to have maximum correlations, can be examined by utilizing canonical loadings. Canonical loadings are the correlations between the original variables and the canonical scores (Stowe vd., 1980: 976).

DATA SETS IN CANONICAL CORRELATION ANALYSIS

Canonical Correlation Analysis is a multivariate method that evaluates the relationship between two sets of variables $(X_1^{(1)}, X_2^{(1)}, ..., X_p^{(1)}; X_1^{(2)}, X_2^{(2)}, ..., X_q^{(2)})$, containing two or more variables by means of linear components. In the Canonical Correlation Analysis, the linear relations between X^1 and X^2 variable sets are investigated. The variable sets have linear components and the correlation between the two sets of variables is calculated with the aid of the canonical variables calculated via the components. The data matrix of X^1 set contains p variables, and the matrix of X^2 data set contains q variables. The variable numbers in the sets must be p> 1, q> 1, and p = q condition is not required (Oktay and Çınar, 2002: 11-31).

With the canonical correlation analysis, new variables are obtained from the linear combinations of the variables in the sets for both variable sets, and it is aimed that the correlation between these new variables is maximum. The new variables obtained are canonical variables, and correlations between them are called canonical correlations (Keskin vd., 2005: 154).

With the Canonical Correlation Analysis, there can be the highest correlations between the linear functions of the two variable sets. In this method, first, each pair of cluster variables has the pair of maximum correlated and unit variance combinations, then the second linear combination pair is obtained and this process is continued (Oktay and Çınar, 2002: 11-31).

If the variable X is expressed as an independent variable and the variable Y is expressed as a dependent variable, that is, if X is interpreted as the reason for Y, V can be called the "best guesser" in this case and "best predictable criterion" (Çankaya vd., 2009: 62).

METHODOLOGICAL APPROACH TO CANONICAL CORRELATION ANALYSIS

It is necessary to decide how many of the canonical variable pairs obtained as a result of the Canonical Correlation Analysis are important, ie, how many of the relationship between the variable groups can be explained in large measure. The purpose of this method is to test how many of the pairs of canonical correlation pairs found are counted as important. In the Wilk's Lambda approach, the hypothesis that all canonical correlations are equal to zero is tested against the alternative hypothesis. The analysis is carried out in the following stages (Oktay and Çınar, 2002: 11-31);

$H_0 : \Sigma_{12} = 0$ or $r_1 = r_2 =$ .... $r_p = 0$

H_A : At least one $r_i \neq 0$

If the H_0 hypothesis is rejected, the largest value will be subtracted from the coefficient hypothesis and the process will be repeated until the H_0 zero hypothesis is accepted. Wilk's Lambda test statistic is obtained as follows.

$$\Lambda = \prod_{i=1}^{k} (1 - r_i^2)$$

Using this coefficient, χ^2_{hes} test statistic value can be calculated as;

$$\chi^2_{hes} = -\left[(n-1) - (p+q+1)/2\right] \log(\Lambda)$$

In this equation n, sample volume; p is the number of variables in the first set; q is the number of variables in the second set; r_i, canonical correlations; k is the number of canonical correlations. The test statistic is compared with the calculated value χ^2_{hes} and $\chi^2_{pxq;\alpha}$ the table value. $\chi^2_{hes} > \chi^2_{pxq;\alpha}$, the H_0 hypothesis is rejected. That is, the first canonical correlation is said to be significant. If the first calculated test statistic χ^2_{hes}, is significant, the first canonical correlation is excluded from the test and the test is repeated with other canonical correlations. This time Wilk's Lambda statistic is calculated for i = 2, 3, ..., k.

$$\Lambda^* = \prod_{i=2}^{k}(1-r_2^2)$$

and

$$\chi^2_{hes} = -\left[(n-1)-(p+q+1)/2\right]\log(\Lambda^*)$$

$> \chi^2_{(p-1)(q-1);\alpha}$

These operations continue until the value χ^2_{hes} is insignificant. Furthermore, as Wilk's Lambertian coefficient approaches zero, the correlation coefficient with the value of H_0 hypothesis rejection (meaning that the canonical correlation coefficient is significant) χ^2 will be different from zero (meaningful).

It is proposed to calculate the redundancy index for each canonical correlation coefficient to determine how much of the total variance in the other set of variables may be explained by a set of variables (Çankaya vd., 2009: 62-63).

If A represents the variation in a set of variables measured over several objects, and B represents variation in another set of measures on the same objects, then A ∩ B may be taken as a measure of their shared variation. This intersection of the two sets is termed "redundancy". Redundancy, expressed as a percentage of the total variation in each set, is rarely symmetrical in the sense that the percentages are the same for both variable sets (Alpert and Peterson, 1972: 188).

The results of the canonical correlation analysis need to be evaluated in terms of linearity, multiple normal distribution, homivariance, and multiple linear connectivity assumptions. The linearity hypothesis affects the results of the analysis of canonical correlation in two ways. First, in canonical correlation analysis it is assumed that the relationship between two variables is linear. If the relationship between the two variables is not linear, then one or both of the variables must be linearly transformed if possible. Second, the canonical correlation coefficients reflect a linear relationship between the two canonical variables. For this reason, nonlinear relations between variables in canonical correlation analysis cannot be explained. In the analysis of canonical correlation, all the metric variables can be used, although a normal distribution assumption is necessary. The reason for the preference of the normal distribution assumption is that the relations between the variables can be increased by standardizing the variables. However, canonical correlation analysis is not affected by normal distribution deviations that do not reduce the relationship between variables (eg, due to distortion of detection from normal distribution). However, in order to be able to test the significance of canonical functions appropriately, it is necessary to use the assumption of multiple normal distribution. Since the multiple normal distribution cannot be easily tested, it is necessary that at least the variables in the analysis are evaluated in terms of the individual normal distribution and the variables that do not correspond to the normal distribution should be converted to the normal distribution if possible. Equivariance should be evaluated separately because it reduces correlations between variables. Finally, it should be ensured that the multicollinearity that makes it difficult to articulate the effects of the variables and thus adversely affects the interpretation of the analysis results to an acceptable level (Kalaycı, 2008: 244-245).

APPLICATION OF CANONICAL CORRELATION ANALYSIS (ANALYSIS OF CANONICAL CORRELATION ANALYSIS BETWEEN TWO DATA SETS COMPOSED OF ECONOMIC AND SOCIAL INDICATORS OF COUNTRIES IN THE ARABIC UNION)

DETERMINATION OF DATA SETS

In the study, the relationship between some choosen economic and social indicators of the 20 countries of the Arab League was analysed. All choosen economic and social indicators for all member countries are determined to use in the analysis.

The data sets obtained for this purpose which are consist of 3 economic and 4 social indicators. The first set of variables, p values, belongs to 20 Arab League member countries; X_1 = Unemployment Rate (Ratio of Total Workforce), X_2 = Total Workforce Quantity (1000 Persons) and X_3 = GNP per Person (Million USD); while the second set of variables q; Y_1 = Population per 1 km², Y_2 = Food Production Index (2004-2006=100), Y_3 = Total Healthcare Expenditures (% GDP) and Y_4 = Total Number of Mobile Phone Subscribers (Per 100 Persons).

The descriptive statistics of the year 2017 data obtained from the official website of the World Bank (URL 1) are presented in Table 1.

Indicators	Minimum Value	Maximum Value	Mean Value	Standart Deviation
Unemployment Rate (Ratio of Total Workforce) (X_1)	0,23	19,22	10,08	5,62
Total Workforce Quantity (1000 Persons) (X_2)	338,08	31569,88	6698,32	7260,07
GNP per Person (Million USD) (X_3)	1727,00	646438,38	122439,67	158246,21
Population per 1 km² (Y_1)	3	1848	185,85	412,73
Food Production Index (2004-2006=100) (Y_2)	62,30	199,40	122,97	34,16
Total Healthcare Expenditures (% GDP) (Y_3)	0,10	10,60	5,20	2,35
Total Number of Mobile Phone Subscribers (Per 100 Persons) (Y_4)	38	217	118,85	51,38

Table 1: Descriptive Statistical Information on Data Used in Canonical Correlation Analysis

DETERMINATION AND ANALYSIS OF RELATIONSHIP BETWEEN ECONOMIC AND SOCIAL INDICATORS

The analysis includes 3 economic indicators and 4 social indicators from 20 Arabic Leauge member countries. Therefore, to analysis the relations between economic and social indicators belonging to countries, it has been examined by analyzing a matrix of dimensions (p + q) x N = (3 + 4) x 20 = 7 x 20.

As a result of the survey, simple correlations between economic and social indicators of the 20 Arabic League member countries are given in Table 2.

	X1	X2	X3	Y1	Y2	Y3
X2	-0,2815					
X3	-0,6453	0,3401				
Y1	-0,4742	-0,1522	0,3283			
Y2	-0,2749	-0,2014	0,3427	0,2949		
Y3	0,2572	-0,0949	-0,4450	-0,0288	-0,0747	
Y4	-0,3517	0,3497	0,6214	0,4509	0,2684	-0,1643

Tablo 2: Simple Correlations between the economic and social indicators of the Arabic League member countries

As a result of the analysis on the correlation coefficients in Table 2;

A correlation of 0.2815 in the reverse direction between the "Unemployment Rate (Total Workforce Ratio)" and "Total Workforce Ratio (1000 Persons)" was found. And also a correlation of 0,6453 in the reverse direction between the "Unemployment Rate (Total Workforce Ratio)" and "GNP per Person (Million USD)" was found for the Arab League member countries. A correlation of 0,3401 was found between "Total Workforce Amount (1000 Persons)" and "Per Capita GDP (Million USD)" for the economic variables of Arabic League member countries.

There was a simple ceorrelation of 0,2949 in the reverse direction between "Population per 1 km²" and "Food Production Index (2004-2006=100)". Accoding to the social indicators, the other simple correlations are; 0,0288 in the reverse direction between "Population per 1 km²" and "Total Health Expenditures (% GDP)" and 0,4509 between "Population per 1 km²" and “Total Number of Mobile Phone Subscribers (100 per Person)” was found.

Likewise, a simple correlation of 0,2747 was found between "Food Production Index (2004-2006 = 100)" and "Total Health Expenditures (% GDP)", and 0,2684 between "Food Production Index (2004-2006 = 100)" and "Total Mobile Telephone Subscribers (Per 100 People)".

Finally, a correlation of 0.1643 was found between "Total Health Expenditures (% GDP)" and "Total Number of Mobile Telephone Subscribers (Per 100 Persons)".

When the correlations of the economic and the social variables in the data sets used in the analysis are examined, a correlation is usually observed in the reverse direction. According to this;

The other variables between the variable "Unemployment Rate"; the simple correlation of 0,4742 in the reverse direction with "Population per 1 km²", 0,2749 in the reverse direction with "Food Production Index (2004-2006=100)", 0,2572 with "Total Health Expenditures (% of GDP)” and 0,3517 in the reverse direction with "Total Number of Mobile Telephone Subscribers (Per 100 People)" was found.

The other variables between the variable "Total Workforce Amount (1000 Persons)"; the simple correlation of 0,1522 in the reverse direction with "Population per 1 km²", 0,2014 in the reverse direction with "Food Production Index (2004-2006=100)", 0,0959 in the reverse direction with "Total Health Expenditures (% of GDP)" and 0,3497 with "Total Number of Mobile Telephone Subscribers (Per 100 People)" was found.

The other variables between the variable GNP per capita (Million USD)"; the simple correlation of 0,3283 with "Population per 1 km²", 0,3427 with "Food Production Index (2004-2006=100)", 0,4450 in the reverse direction with "Total Health Expenditures (% of GDP)" and 0,6214 with "Total Number of Mobile Telephone Subscribers (Per 100 People)" was found.

The results based on the relationship between canonical variables are shown in the Table 3.

Canonical Variables (U_1, V_1)	**Canonical Correlation**	**R-Square Value**	**Wilk's**	**Chi-Square**	**Degree of Freedom**	**Significance Level**
1.	0,734	0,538	0,242	21,353	12	0,047
2.	0,670	0,449	0,52	9,637	6	0,141
3.	0,212	0,045	0,955	0,690	2	0,708

Tablo 3: Relationship Results Between Canonical Variables

As a result of the examination made on Table 3;

- The canonical correlation coefficient between U1 and V1 canonical variable pair was calculated as ρ1 = 0,734 and p < 0,05 (p = 0,047) and statistically significant,
- The canonical correlation coefficient between U2 and V2 canonical variable pair was calculated as ρ2 = 0,670 and p > 0,05 (p = 0,141) and statistically insignificant,
- The canonical correlation coefficient between U3 and V3 canonical variable pair was calculated as ρ3 = 0,212 and p > 0,05 (p = 0,708) and statistically insignificant.

At this stage of the method, only significant coefficients can be interpreted. The first canonical variable pair with significant and biggest canonical correlation will be considered in the analysis. According to the first canonical variable pair, it can be said that the independent variable set can explain the dependent variable set at 54% and the rest is under the influence of other factors.

The canonical weights and canonical functions calculated for the first set of independent and dependent variables, which are determined to be significant and have the highest correlation with respect to the relationship between them, can be expressed as follows:

$$U1 = -0,036\ x1 + 0,045\ x2 - 1,037\ x3$$

$$V1 = -0,062\ y1 - 0,250\ y2 + 0,480\ y3 - 0,665\ y4$$

The extent to which the original variables contribute to the canonical variance can be seen from the functional equality given above. Accordingly, the original variables explaining the first canonical variables of economic data belonging to the Arabic League member countries are GDP per capita (USD million), Total Workforce Amount (thousand persons) and Unemployment Rate (Ratio to Total Workforce). In the same way, the original variables that explain the first canonical variable belonging to the social data are the total number of Mobile Phone Subscribers (per 100 persons), Total Health Expenditures (% GDP), Food Production Index (2004-2006 = 100).

Canonical load is expressed as a simple linear correlation between the original variable and its canonical variant.

Canonical load is determine how strong the contribution of the respective variable is to the canonical correlation coefficient of its canonical variable. Also canonical cross-load is defined as simple linear correlation between original dependent variables and independent canonical variables or simple linear correlation between original independent variables and dependent canonical variables. Thus, the power of the observed variation with high correlation can be measured by the canonical variation in the cross-set (Bilgin et al., 2003: 345-346). Table 4 below shows the canonical loads and cross-loads of economic data included as an independent variable in the analysis. Similarly, canonical loads and cross-loads of social data are presented in Table 5.

Tablo 4: Canonical Loads and Canonical Cross-loads of Economic Variables

	U1	**V1**
x1	0,621	0,456
x2	-0,298	-0,219
x3	-0,999	-0,733

Tablo 5: Canonical Loads and Canonical Cross-loads of Social Variables

	U1	**V1**
y1	-0,450	-0,330
y2	-0,483	-0,355
y3	0,610	0,488
y4	-0,839	-0,616

From Table 5, it can be seen that the variable having the highest simple linear correlation coefficient with dependent and independent canonical variable is “Unemployment Rate (Ratio to Total Labor)” from independent original variables in the study. That is, the highest contribution to dependent and independent canonical variables was “Unemployment Rate (Ratio to Total Labor).

Likewise, as can be seen in Table 6, it is determined that “Total Number of Mobile Phone Subscribers (per 100 persons)” is the variable having the highest simple linear correlation coefficient with dependent and independent canonical variable among social variables. Here again, it can be stated that the dependent variable and independent canonical variables provide the highest contribution by this original dependent variable.

CONCLUSION

The Arab League (al-Jāmiʻah al-ʻArabiyah), formally the League of Arab States (Jamiʻat ad-Duwal al-ʻArabiyah), is a regional organization of Arab states in and around North Africa, the Horn of Africa and Arabia. It was formed in Cairo on 22 March 1945 with six members: Kingdom of Egypt, Kingdom of Iraq, Transjordon (renamed Jordın in 1949), Lebanon, Saudi Arabia, and Syria. Yemen joined as a member on 5 May 1945.

The other members are Libya (1953), Sudan (1956), Tunisia and Morocco (1958), Kuwait (1961), Algeria (1962), Bahrain, Oman, Qatar and United Arab Emirates (1971), Mauritania Somali (1974), Palestinian Liberation Organization (PLO, 1976) and Djibouti (1977). (Both Yemen were members of the organization between 1967-1990.) In 1993, the number of members increased to 24 with the membership of Komor Islands.

In this study, the relationship between some economic and social indicators on the basis of countries was determined with the help of canonical correlation analysis, 20 Arabic League countries which can fully access the data used in the analysis.

According to the aim of this study, the economic variables obtained from the World Bank Official Web Site for 20 Arabic League member countries; X_1 = Unemployment Rate (Ratio to Total Labor), X_2 = Total Workforce (1000 People) and X_3 = GNP per Person (Million USD). Similarly, the second set of variables obtained as dependent variables is Y_1 = Population per 1 km², Y_2 = Food Production Index (2004-2006=100), Y_3 = Total Healthcare Expenditures (% GDP) and Y_4 = Total Mobile Phone Subscribers (per 100 Person) was used in the analysis.

Only one variable pair with the highest value among canonical variables was found statistically significant. As a result of examining the values in Table 4, it was seen that the U_1 and V_1 couples gave meaningful results, and that the couples had the greatest contribution to the total variance. It can be said that the other variable pairs are statistically insignificant because they do not contribute too much to the total variance.

The "declared variance ratio", which indicates the part of the canonical variables described in their set; is calculated by taking the average of the squares of the canonical loads of each canonical variable in the dependent or independent set. The variance ratio of the canonical variables obtained from the dependent set is 100% in total. The variance ratio of the first dependent canonical variable (V_1) is 37.8%. The variance ratio of the canonical variables obtained from the independent set in its set is 100% in total. The variance rate of the first independent canonical variant (U_1) is 49.0%.

The redundancy index refers to the part of the canonical variables described in the cross-set, and i^{th} with the explained variance ratio of the canonical variable i^{th} canonical correlation coefficient. According to the redundancy index, the independent canonical variables explained 38.2% in the dependent set and the dependent variable was 29.6% in the independent set. The fraction of the first dependent canonical variable explained in the independent set is 20.4%. Likewise, the first independent canonical variable explained in the dependent set is 26.4%.

Correlation coefficients and canonical correlation coefficients calculated between variables and U1:V1 calculated as ρ1 = 0,734 for canonical variable pair, ρ2 = 0,670 for U2:V2 canonical variable pair and ρ3 = 0,212 for U3:V3 canonical variable pair were found.

This can be reiterated in other studies by adding more variable groups and/or other countries for this situation which is assessed as being caused by the number of Arabic League member countries.

REFERENCES

1. ALPERT M.I. and Peterson R.A. (1972). "On The Interpretation of Canonical Analysis", Journal of Marketing Research, Vol. IX.
2. BAGOZZI R.P., Fornell C. and Larcker D.F. (1981). "Canonical Correlation Analysis as a Special Case of a Structural Relations Model", Multivariate Behavioral Research, 16.
3. ÇANKAYA S., Altop A., Oflaz M. and Erener G. (2009). "Karayaka Toklularında Kesim Öncesi ve Kesim Sonrası Ölçülen Bazı Özellikler Arasındaki İlişkinin Tahmini İçin Kanonik Korelasyon Analizi", Anadolu Tarım Bilimleri Dergisi, 24 (1).
4. FORNELL C. and Larcker D.F. (1980). "The Use of Canonical Correlation Analysis in Accounting Research", Journal of Business Finance & Accounting, 7, 3.
5. GARCIA-Gallego A. and Mures-Quintina M. (2016). "Principal Components and Canonical Correlation Analyses as Complementary Tools. Applicatıon to the Processing of Financial Information", Economic Computation and Economic Cybernetics Studies and Research, Issue 4, Vol. 50.
6. KALAYCI Ş. (2008). SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri, Asil Yayın Dağıtım Ltd.Şti, Ankara.
7. KESKİN S., Kor A. and Başpınar E. (2005). "Akkeçi Oğlaklarında Kesim Öncesi ve Kesim Sonrası Ölçülen Bazı Özellikler Arasındaki İlişki Yapısının Kanonik Korelasyon Analizi ile İrdelenmesi", Tarım Bilimleri Dergisi, 11 (2).
8. OKTAY E. and Çınar H. (2002). "Avrupa Birliği Ülkelerinin Bazı Ekonomik ve sosyal Göstergeleri Arasındaki İlişkinin Kanonik Korelasyon Analizleri Yardımıyla Belirlenmesi", EKEV Akademi Dergisi, Yıl:6, Sayı:12.
9. PUGH R.C. and Hu Y. (1991). "Use and Interpretation of Canonical Correlation Analysis in Journal of Educational Research Articles: 1978-1989", Journal of Educational Research, Vol.: 84, No.: 3.
10. STOWE J.D., Watson C.J. and Robertson T.D. (1980). "Relationships Between the Two Sides of the Balance Sheet: A Canonical Correlation Analysis", The Journal of Finance, Vol. XXXV, No. 4.
11. URL 1: https://data.worldbank.org/ indicator, Date of Access: 08 May 2018.