مجلة التجارة والتمويل

/https://caf.journals.ekb.eg

كلية التجارة – جامعة طنطا

العدد : الثانى

يونيو ٢٠٢٢

An Application of Linear Programming Discriminated Analysis for Classification

Maie M., Kamel , Hanaa A. Salem
and Waleed Abdelgawad.
Faculty of commerce, Tanta University
Statistics Department
maie.kamel@commerce.tanta.edu.eg
hanaa_salem@commerce.tanta.edu.eg
walidabdelgwad2@gmail.com

أ.م.د/ مي محمد كامل
استاذ مساعد الاحصاء كلية التجارة جامعة طنطا
أ.م.د/ هناء عبد الرحيم سالم
استاذ مساعد الاحصاء كلية التجارة جامعة طنطا
وليد عبد الجواد
تاهيلي دكتوراه قسم الاحصاء كلية التجارة جامعة طنطا

An Application of Linear Programming Discriminated Analysis for Classification

Abstract

The goal of this study is to compare linear discrimination analysis and discriminated analysis with linear programming (MMD) (Min. Sum of Deviation) in order to find the best model for classifying observations into their correct groups with the lowest possible classification error and highest classification accuracy. According to the findings of the study, discriminated analysis using linear programming differs from linear discriminated analysis in data classification because it produces the lowest error rate and the highest classification accuracy rate, and it does not require the linear discriminated analysis assumptions.

keyword: Discriminated analysis, discriminated analysis using mathematical programming, corporate bankruptcy, data classification

Introduction

There are a variety of statistical methods used in the analysis of metadata, as the method used in this analysis is dependent on the ability to classify the variables into dependent and independent variables, as well as the number and quality of these variables, whether quantitative or qualitative data (Thabit et al. 2009).

T. Tony Cai and Linjun Zhang (2019) established optimality theory for high-dimensional linear discriminated analysis. An adaptive restricted l1-minimization strategy was used to propose and analyse a data-driven and tuning-free classification rule. When data is high-dimensional and local aspects of the data might play key roles in classification, Jiang, Binyan, Ziqi Chen, and Chenlei Leng (2020) suggested a new and simple model for high-dimensional linear discriminated analysis. For analyzing high-dimensional problems, we combine the simplicity of kernel smoothing with the strong method of regularization.

Wasef, Michael R., and Nader Rafla (2020) explained the classifier design and optimization method and provided data on power consumption, resource utilization, latency, and the algorithm accuracy before and after optimization.

There are some methods that do not distinguish between dependent and independent variables in a study, and the most important of these methods is factor analysis, which is used in the field of applied statistics to reduce the number of variables from a large number to a small number of factors based on the highest value of the correlation between the variables. This method implies that research variables be quantitative, and cluster analysis, which collects observations with similar characteristics and organizes them into clusters so that classified instances within each cluster are homogeneous, is regarded one of the most prominent descriptive methods.

We can't define sectors ahead of time in cluster analysis. Where analysis is done primarily to identify those sectors, as opposed to some other statistical methods that allow us to pre-define sectors, such as Timm (2002). There is another multivariate statistical method called Discriminated Analysis, which is used in a variety of applications, including economics, social science, medicine, business administration, and education.

The following goals are the focus of this study:

1.Theoretical underpinning for each method of analysis of variance with linear programming is presented.

2. Developing a linear discriminated function that may be used to classify enterprises as bankrupt or non-bankrupt based on a set of criteria (the ratio of cash flow to total liabilities, the return on assets, the current asset-to-sales ratio, the return on equity).
3.Creating a linear programming model that will be used to categorise fresh views into one of the two groups (bankrupt companies and non-bankrupt companies).

4.Compare the two systems and determine which is more accurate in terms of categorisation

The purpose of this research is to compare the classification accuracy of linear discriminated analysis versus discriminated analysis with linear programming. The research's limitations are presented in the second section, and the theoretical background for each of the two methodologies is presented in the third section. The findings of two methods, as well as the recommendations, are presented in the fourth section..

Search Limits:
Data were collected through the financial statements for the year 2017, which is available on the Internet through the following website: https://www.mubasher.info/countries/eg/ companies

Based on the Altman model, the companies were separated into bankrupt and non-bankrupt categories, and a sample of 36 companies was chosen for data availability and divided into two equal groups. There are 18 bankrupt companies in the first category, and 18 non-bankrupt enterprises in the second.

Research problem

The linear discriminated analysis method, despite being the most widely used method for categorizing data, has certain drawbacks, the most significant of which are: (Mustafa et al., 2014)

1-The issue of the small sample sizes, where the size of the training sample may be small in comparison to the size of the community from which it is extracted.

2-One of the hypotheses for linear discriminated analysis is contradicted by the heterogeneity of the variance and covariance matrix of random variables (Covariance and covariance matrix is equal) $\sum_1=\sum_2$

To solve these issues and the inability of linear discriminated analysis to achieve one or more hypotheses, researchers have been turned to discriminated analysis utilizing LPDA mathematical programming.

Discriminated analysis

Fisher discovered the linear discriminated function in 1936 and used it to classify two groupings of plants. Because of the concept of linear discriminated analysis, the differences between the two groups are substantially bigger than the differences within them. The goal was to classify the observation into the appropriate category and using this method, the degree of classification reaches its highest level in terms of achieving the goal. The following are the underlying assumptions:

1- That the data is distributed according to a multivariate normal distribution.

2- Variance and Covariance matrices of the statistical societies under investigation are equal. $\sum_1=\sum_2$

3-Independence of observations.

The goal of the discriminated analysis is to discriminate between observations from various and preset societies, as well as to anticipate whether a new observation belongs to one of the groups. The approach of discriminated analysis is based on obtaining a Discriminated function, which is a linear combination of independent variables. This function increases the differences between the mean Groups, where the larger the spacing between the mean of the groups, the more effective the distinction and consequently the lower the classification mistakes.

The following are the main assumptions of discriminated analysis:

1- Analysis of variance is based on a set of assumptions that are similar to those used in multivariate analyses like multivariate analysis of variance, namely: The observations are distributed in a multivariate normal distribution.

2- The homogeneity of the covariance and covariance matrix of the study population$\sum_1 = \sum_2$views independence

3- Characteristics of discriminated analysis:

The discriminated analysis approach has the following main characteristics:

1- It is regarded as one of the most advanced statistical methods for multivariate data analysis.

2 -Because prior information of the sectors is required for analysis of variance, discriminated analysis varies from other statistical procedures (cluster analysis)

3- In discriminated analysis, societies are classified based on a value known as the cut point, which provides a single number from which we can assign each individual to its original community.

4-In terms of error rates, the discriminated analysis results are more accurate than the logistic regression results.

The discriminated analysis approach is a descriptive method that may be used to explain what happened in the past and predict what will happen in the future (Thabit, others 2009).

Objectives of the method of discriminated analysis:

The following are the goals of discriminated analysis:

1- Differentiate between observations from diverse and preset societies, as well as forecast which of the groups (societies) under study any new observation belongs to.

2- Identifying the most influential independent variables that can be used to classify the dependent variable.

3- Investigate the differences between two groups based on a collection of independent factors in order to distinguish them and categorize observations into the appropriate category.

4- Calculate the proportion of correct classifications.

The efficiency of the linear characteristic function is put to the test.

This is one of the most significant steps in the discriminated analysis since it determines the function's ability to differentiate. The F test is used to discriminate between the two groups. We can utilize this when we need to distinguish between two groups. The equality of the two groups is tested..

$H_\circ$: $\mu_1 = \mu_2$ The function does not have the ability to distinguish

H_1:$\mu_1 \neq \mu_2$ The function has the ability to distinguish

The Wilkes scale, which is calculated using the following formula, can be used to measure the effectiveness of the linear characteristic function: (Al-Mikhlafi, 2014)

$$\Lambda = \frac{|G|}{|T|}$$

Where G: the scattering matrix within groups.

T: total dispersion matrix.

The value of (Λ) ranges from zero to one. If its value is close to zero, then this indicates the strength of the discrimination and x^2 can be calculated from the scale and its formula is as follows:

$x^2 = -\log(\Lambda)$
With a degree of freedom P (K - 1) where P is the number of variables, K is the number of groups
Homogeneity test between groups using the variance matrices
This test is intended to select the best model to represent the discriminated function and have a high classification capacity.
And the statistical assumptions are as follows:

$$\boldsymbol{H_0}: \sum\nolimits_{\mathbf{1}} = \sum\nolimits_{\mathbf{2}} = \cdots = \sum\nolimits_{\mathbf{k}}$$

$\boldsymbol{H_1}$: At least one is different.

Box's-M test takes the following form:

$$M = \begin{cases} (n-k)\log|S| - \sum_{i=1}^{k}(n_i - 1)log|S| & if \ |S| > 0 \\ SYSMIS & if \ |S| \le 0 \end{cases}$$

$$S = \begin{cases} \frac{1}{(n-k)}\sum_{i=1}^{k} (n_i - 1)\, S_i & if \ n > k \\ 0 & if \ n < k \end{cases}$$

Where S the combined covariance matrix
Si: the covariance matrix of the group where i=1,2,,..,k
n_i the number of group views
n total sample size
r : number of independent variables
k: number of groups
p: length of the column vector

the test F (γM, f_1, f_2) is used to calculate the significance, where

$$f_1 = (k-1)\,p\,(p+1)/2$$

$$r = 1 - \frac{2p^2 + 3p - 1}{6\,(p+1)(k-1)}\left[\sum_{i=1}^{k}\frac{1}{(n_i - 1)} - \frac{1}{(n-k)}\right]$$

$$\tau = \frac{(p-1)(p+2)}{6(k-1)}\left[\sum_{i=1}^{k}\frac{1}{(n_i - 1)^2} - \frac{1}{(n-k)^2}\right]$$

$$f_2 = \frac{(f_1 + 2)}{|\tau - (1-r)^2|}$$

$$\gamma = \frac{(r - \frac{f_1}{f_2})}{f_1}$$

Where

The value of F is then compared to the Box's-M statistic. We reject the null hypothesis and accept the alternative hypothesis if the tabular value of F is smaller than the statistic of Box's-M, i.e. we reject the hypothesis of homogeneity of variance. When the sample size is small or average, this test is preferred.

The Bartlett test can be used to find the best sort of model to use to represent the discriminated function in big samples.

The following are the statistical assumptions:

$$\boldsymbol{H_0}: \sum_{\mathbf{1}} = \sum_{\mathbf{2}} = \cdots = \sum_{\mathbf{k}}$$

$$\boldsymbol{H_1}: \textbf{At least one is different}.$$

The format of this test is as (Talib et al. 2019):
follows

$$M = In\,|S|\,(\textstyle\sum_{i=1}^{k} V_i\,) - \textstyle\sum_{i=1}^{k}\,(V_i\;In\,|S|\,)$$

$$,\, S_p = \frac{1}{\sum_{i=1}^{k} V_i}\sum_{i=1}^{k} V_i\,S_i$$

Where : S_p is an estimate of the covariance matrix.

S_i is an estimate of sample variance where i= 1,2,3,…,k

$$V_i = n_i - 1$$

In 1949, Box proved that when multiply M in a constant C^{-1} which equal

$$C^{-1} = 1 - \frac{2p^2 + 3p - 1}{6\,(p+1)(k-1)}\left[\sum_{i=1}^{k}\frac{1}{V_i} - \frac{1}{\sum_{i=1}^{k} V_i}\right]$$

will have a scale distributed to chi square whose degrees of freedom equal to $\left[\frac{1}{2}(k-1)p\,(\,p+1)\right]$

M $c^{-1}\sim X^2\left[\frac{1}{2}(k-1)p\,(\,p+1)\right]$

where :

p is no of independent variables, K is no of groups

Then the expression ((M c^{-1})) is compared with the tabular value of χ^2, and if the tabular value of x^2 is less than the amount ((M c^{-1})) then we reject the null hypothesis and accept the alternative hypothesis, i.e. we reject the hypothesis of homogeneity of variance.

Discriminated analysis using mathematical programming

Mathematical programming is a mathematical method for determining the greatest or minimum value of a certain function called the objective function that is dependent on a set of variables. It is one of the most important models of operations research that is utilized in decision-making. These factors may be independent or connected in what is referred to as constraints.
There are numerous advantages of using mathematical programming in discriminated analysis (Zeidan 2012)
1 - It can be used without making any assumptions.
2 - Mathematical programming can benefit from sensitivity analysis.
3 - In mathematical programming, determining individual weights (coefficients of discriminated) for all variables is simple.
4- Unlike traditional methods of discriminated analysis, mathematical programming methods are less subject to external observations.
However, one of the disadvantages of LPDA mathematical programming models are that some of them, particularly Goal Programming (GP) models, have infinite solution values.
There are a variety of linear programming models that can be used to address the discriminated analysis problem, and they are listed below:
The MSD -1 model (Minimal Sum of Deviation) /
Freed and Glover (1986) provided a mathematical programming model, which was used to develop this model based on the method of lowering the sum of the external deviations of the observations from the group to which they should belong. The formula for the model is as follows: (Alsaid, Alfraidy,2017)

$$\min\left(\sum_{i=1}^{p} d_i\right)$$

Subject to :

$$\sum_{j=1}^{p}(w_j x_{ij} + d_i) \geq c \qquad , i \in G_1$$

$$\sum_{j=1}^{p}(w_j x_{ij} - d_i) \leq c \qquad , i \in G_2$$

$\sum_{j=1}^{p} w_j$=1(Normalization constraint)

$d_i \geq 0 , (i = 1,2, \ldots , n)$

Where we are required to find the values of w, d, c

d_i: It represents the sum of the external deviations of which i from its group to which it should belong, and the external deviation is defined as the distance between the wrongly rated watch and the boundary of its group

x_{ij} Indicates the value of the variable j for observation i,
where i = 1,2, ..., n, j = 1,2, ..., p

c: which is the boundary of the two groups (the breakpoint).

w_j Denotes the coefficients of the discriminated functions.
where c, w is not specified

constraint$\sum_{j=1}^{p} w_j$means a zero solution prevents
Each item (i) is classified into its group to which it should belong according to the value of the discriminated function. If the value of the discriminated function is greater than the constant c (the cutoff point), then the observation belongs to a specific group and if the value of the discriminated function is less than this constant, then it belongs to the group the second where c ,w_j unrestricted in sign.

The applied study

The application was carried out on actual data of a group of companies for the purpose of predicting financial failure. The SPSS program was used to obtain the results of discriminated analysis and the WinQSP program to obtain the results of the discriminated analysis using linear programming.
In the applied study, Y was: it is a binary dependent variable that takes the value (1) if the watch i belongs to the first group (companies are not bankrupt), the number (2) if the watch i belongs to the second group (companies are bankrupt)

$$Y_{ki} = \begin{cases} 1 & if\ observation\ i\ belongs\ to\ first\ group \\ 2 & if\ observation\ i\ belongs\ to\ second\ group \end{cases}$$

Where i = 1 , 2 ,…….n , k = 1, 2,3…

The independent variables are

Ratio of cash flow to total liabilities $X_1 = \frac{cash\ flow}{\text{total liablity}}$

Return on assets $X_2 = \frac{net\ income}{\text{total assets}}$

Return on equity$X_3 = \frac{net\ income}{\text{ownes equity}}$:

The ratio of current assets to sales$X_4 = \frac{current\ assets}{\text{sales}}$

Verification of assumptions about analysis of discriminated:
Test the normal distribution of the data
To verify this hypothesis, the kolomgrove Saminov test is used, and this is done through the following statistical hypothesis test:
null hypothesis H_0: Data follow a normal distribution
Alternative Hypothesis H_1 : Data do not follow a normal distribution

Table(1)

kolomgrove Smirnov test on normal distribution of data

Decision	Sig	Variable
Retain the null hypothesis	. 270	X_1
Retain the null hypothesis	. 270	X_4
Retain the null hypothesis	.057	X_3
Reject the null hypothesis	.001	X_2

. The significance level is .05

From the results of Table (1), it was found that the ariables X_(1) X_(3) X_ (4) follow the normal distribution, depending on thve kolomgrove Smirnov , where it becomes clear that the value of Sig for those variables is greater than 0.05 except for the variable X_2 and according to the central tendency theory This data follows the normal distribution because data size is greater than 30 views.

The discriminated function significance test: This step is considered one of the most important steps in discriminatory analysis To test the ability of the function to distinguish, Wilk's Lambda test is used, and this is done through the following statistical hypothesis test:

$H_\circ: \boldsymbol{\mu_1} = \boldsymbol{\mu_2}$

$H_1: \boldsymbol{\mu_1} \neq \boldsymbol{\mu_2}$

Assumptions can be formulated in another way

H_0: The function does not have the ability to distinguish

H_1: The function has the ability to distinguish

Table(2)

The discriminated function significance test

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig
1	0.545	19.411	4	0.001

Through the results listed in Table (2), we notice that the significant value is equal to 0.001 and it is less than 0.05 which indicates that there are significant differences between the bankrupt companies group and the non-bankrupt group of companies, and this means that the discriminatory function has the ability to distinguish and classify observations to their real society.

Test for the homogeneity of variance between the two groups:

One of the hypotheses for analysis of discriminated is the assumption of the homogeneity of the variance matrix between the two groups. The Box'M test is used to verify this hypothesis and this is done through the following statistical hypothesis test:

$$H_0: \sum_1 = \sum_2$$
$$H_1: \sum_1 \neq \sum_2$$

To test the homogeneity of the variance between the two groups, the Box'M test is calculated

The following table shows the results of the Box'M test

Table(3)

The homogeneity of the contrast test between the two- group

Box's M		86.384
F	Approx**.**	7.530
	significant.	0**.**000

Tests null hypothesis of equal population covariance matrices

The results in Table (3) indicate that the value of the Box'M statistic is equal to 86.384 and it is noticed that the significant value is equal to 0.00 and it is less than 5%, which indicates the rejection of the null hypothesis and the acceptance of the alternative hypothesis, i.e. the heterogeneity of the disparity between the two groups, and thus the condition of the linear discriminated function will not be fulfilled.

Relationship Strength Test

Table(4)

The Function of Eigen values

Function	Eigen value	% of Variance	Cumulate%	Canonical Correlation
1	0.834	100.0	100.0	0.674

- From Table (5-4), the following results were shown:
- That the correlation coefficient is equal to 0.674, meaning that there is an intermediate relationship between the dependent variable and the independent variables.
- The distinctive value ((Eigen values) equal to 0.834, which indicates the ability of the function to distinguish, as the more the distinctive value approaches the correct one, indicating the ability of the function to distinguish, and what confirms that result is that 100% of the variance was explained.
- Classification results for discriminated analysis

Table(5)

The Results of correct and incorrect classification of discriminated analysis

		Predicted Group Membership		
	Type	**Non** Bankrupt companies	Bankrupt companies	**Total**
Original Count	Non Bankrupt companies	16	2	18
	Bankrupt companies	5	13	18
%	Non Bankrupt companies	88.9	11.1	100.0
	Bankrupt companies	27.8	72.2	100.0

80.6% of original grouped cases correctly classified

Table (5) shows the results of the classification for discriminated analysis of 36 bankrupt and non-bankrupt companies. The first group (non-bankrupt companies) of 18 companies was classified according to the following classification rule:

 16 companies were classified into the first group (non-bankruptcy companies) out of 18 companies with a correct distinction of 88.9%

 As for the second group, a group of 18 bankrupt companies, according to the following classification rule:

Thirteen companies were classified into the second group (bankrupt companies) out of 18 companies with a correct distinction rate of 72.2%, and the average overall correct rating rate is 80.6%.

A- The apparent error rate in the first community (non-bankrupt companies) is as follows:

$$\hat{P}_{21} = \frac{n_{12}}{n_1} = \frac{2}{18} = 11\%$$

B- The apparent error rate in the second community (bankrupt companies) is as follows:

$$\hat{P}_{12} = \frac{n_{21}}{n_2} = \frac{5}{18} = 27\%$$

Classification results using LPDA mathematical discriminated analysis

The solution form was used WinQSB Quantitative System for Business plusLP MSD 41 var 37 const

By applying the model, it was found that the value of the objective function was reduced to 0.0861,the value of the cutoff point equal to 0.0341, and the value of the discriminated coefficients as follows:
w_1 =0.0559 ,$w_2 = 0.8815$, $w_3 = 0.0672$, w_4= -0.0045
and the value of the deviations is as follows:
From the group that 11 means the external deviation in the observation number d_11 = 0.0187 0.0187 belongs to it
For the group that should 15 means the external deviation in the observation number d_15=0.0020
For the group that must be17 means the external deviation in the observation number d_17=0.0089
For the group that must be 21 means the external deviation in the observation number d_21=0.0313
D_25=0.0236 means the external deviation in the observation number 25 from the group it should belong to 0.0236
D_31=0.0016 means the external deviation of the observation number 31 from the group it should belong to 0.0016

A table(6) showing the correct classification percentage for both bankrupt and non-bankruptcy companies Using the MSD model

Correct Rating Ratio%	Non-bankrupt companies	Bankrupt companies	The company
100	0	18	Bankrupt companies
100	18	0	Non-bankrupt companies

It turns out that the overall correct rating percentage is 100%

First: linear programming models
In the MSD-1 model we get the minimized value of objective function is 0.0861 and the cut point value is 0.0341 .The value of the discriminate coefficients are

w_1 = .0559 ,w_2 = .8815 ,w_3 = .0672, w_4= -.0045

Table(7)
Table showing classification results using linear programming models

Model	The first group (Non-bankruptcy companies)		the second group (Bankrupt companies)		Total views rated correctly	Match rate%
	The number of views categorized correctly	The number of categorized views is an error rating	The number of views categorized correctly	The number of categorized views is an error rating		
MSD-1	18	0	18	0	36	100

Second: The results of classification using the Fisher function

Table (8)
Table showing the classification results using the Fisher function

Model	The first group (Non-bankruptcy companies)		the second group (Bankrupt companies)		Total views rated correctly	Match rate%
	The number of views categorized correctly	The number of categorized views is an error rating	The number of views categorized correctly	The number of categorized views is an error rating		
FLDF	16	2	13	5	29	80.6

When applying the linear discriminated analysis method, it was found that the overall correct classification ratio reached 80.6%, which is somewhat high, .and therefore the error classification rate was 19.4%
When comparing the discriminated analysis using linear programming, the practical application proved that it achieved better results than linear discriminated analysis in terms of classification accuracy

Recommendations

Based on the findings of the study and the foregoing, the researcher can make the following recommendations:

1- The use of operation research models as one of the most important and best methods for classifying companies into bankrupt and non-bankrupt companies, particularly companies that are not listed on the stock exchange, because one of the most difficult challenges facing financial analysts is the inability to apply some of the models used in classification because they require data on the market value of the share.

2- Finding the best of the other differential functions (mixed, quadratic, and logistic) in terms of classification accuracy using the other differential functions (mixed, quadratic, and logistic).

2- Classifying companies into bankrupt and non-bankrupt using indications other than financial data.
3- When classifying data, using different operations research models (target programming).

5- The application of sensitivity analysis, which allows researchers to investigate changes in the values of model variables (the ratio of cash flow to total liabilities, the ratio of current assets to sales, the return on equity, and the return on assets) and their impact on model results. company performance, particularly in light of the current economic climate

6- The importance of using mathematical programming models to the Egyptian market in order to forecast current financial situations.

Non-bankrupt companies	**X4**	**X3**	**X2**	**X1**
Egypt BeniSuef Cement	0.9568	0.0618	0.0436	-0.0732
Sir Kraer	0.5915	0.0201	0.230	0.0284
Kafr El Zayat for pesticides	0.8008	0.2108	0.0921	0.3282
Arab Medicines	0.5512	0.2264	0.1348	0.1307
Mina Pharma	0.9471	0.1042	0.0395	0.0098
October Pharma	0.7540	0.3610	0.1518	-0.1342
North Cairo Mills	0.1175	0.2647	0.1372	0.0596
Arab Cotton Ginning	0.8530	0.0944	0.0704	-0.1372
Heliopolis Housing	2.116	0.5043	0.1728	-0.1133
Foodico	0.5377	0.1627	0.0912	-0.0206
Misr Oils and Soap	0.3892	0.1236	0.0098	0.0040
Delta Sugar	0.5239	0.3217	0.2150	-0.3075
Mansoura Poultry	0.8121	0.1067	0.0944	-1.185
Eastern Company	1.175	0.2308	0.0982	0.4790
Oriental carpet weavers	2.169	0.0409	0.0282	0.2573
United Housing	3.395	0.2505	0.1834	0.8336
National	2.185	0.0462	0.0414	-0.0792

Housing				
The Sun for Housing	3.105	0.2888	0.1309	0.0650
bankrupt companies	**X4**	**X3**	**X2**	**X1**
Egyptian iron and steel	1.542	0.3386	-0.2203	0.0188
Sinai Cement Company	0.5305	-0.3679	-0.1589	0.0150
Arab Cement	0.1780	0.1660	0.0621	-0.0060
Tora Cement	0.1328	2.457	-0.7036	-0.0340
Egypt Cement Qena	7.026	0.0201	0.0087	0.2180
Telecom	0.4882	0.4713	0.0078	-0.0388
Telecom Egypt	0.4978	0.0227	0.0122	-0.0223
Pera Misa for hotels	2.576	0.0980	0.0673	0.0635
Golden Pyramids Plaza	2.934	0.0525	0.0268	-0.0091
Nile for Medicines	1.0423	0.0729	0.0281	-0.0246
Egypt Gas Company	0.9990	0.2709	0.0077	-0.0261
Cairo Housing	1.865	0.0297	0.0115	-0.0060
Zahraa El Maadi	11.79	0.0217	0.1013	-0.0527
El Obour Real Estate Investment	2.294	0.0578	0.0187	-0.1311
Egyptian Starch Industry	0.3098	0.0704	0.0370	-0.0042
KomOmbo Valley	2.703	-0.0850	0.0398	0.0002
Orange	5.219	-0.1352	-0.0599	0.0031
North Upper Egypt to develop agricultural production	2.439	-0.0680	-0.2746	-0.5215

References:

Allandale, Abdulkader ,2017,“ **Multivariate Methods and Forecasting with IBM , SPSS Statistics ",**ISBN-13: 978-3319564807,ISBN-10: 3319564803

[1] Bal, H. , Örkcü, H. H. &Celebioğlu, S.,2006, " **An Experimental Comparison of The New Goal Programming and The Linear programming Approaches in The Two-Group Discriminated Problems** " ,Computers and Industrial Engineering , Vol.50, pp. 296-311,https://doi.org/10.1016/j.cie.2006.06.002.

[2] Ben Youssef, S. &Rebai, A., 2007,**" Comparison between Statistical Approaches and Linear Programming for Resolving Classification Problem** " International Mathematical Forum , Vol.2 , No. 63 , pp. 3125-3141,doi:10.12988/IMF.2007.07288 CorpusID:17830805.

[3] Cai, T. T. and Zhang, L. ,2019,” **High dimensional linear discriminated analysis: optimality, adaptive algorithm and missing data**”, J. R. Statist. Soc. B , 81, Part 4, pp. 675–705,arXiv:1804.03018v1.

[4] Glen,J.J., 2001,“ **Classification Accuracy in Discriminated Analysis**" The Journal of operational research society ,Vol. 52 , No. 3,https://doi.org/10.1057/palgrave.jors.2601085 .

[5] Glen, J. J., 2005, “**Mathematical Programming Models for Piecewise -Linear Discriminated Analysis** "university of Edinburgh,.https://www.jstor.org/stable/4102132 .

[6] Härdle ,W., Simar, L., 2015,"**Applied Multivariate Statistical Analysis**",4th edition,DOI 10.1007/978-3-662-45171-7 .

[7] **Holden, J. E., Kelly, K.,2010," The Effects of Initially Misclassified Data on the Effectiveness of Discriminated Function Analysis and Finite Mixture Modeling"** , Education and Psychological Measurement ,https://doi.org/10.1177/0013164409344533

[8] Jiang, Binyan, Ziqi Chen, and Chenlei Leng. ,2020,"Dynamic linear discriminated analysis in high dimensional space." *Bernoulli* 26.2 p.p.1234-1268.

[9] Kwak, N. K.,Kim,S.,H., Lee,C.,W., Choi,T.,S.,2002,“ **An Application of Linear Programming Discriminated Analysis to Classifying and Predicting The Symptomatic Status of HIV \ AIDs Patients** “Journal of medical systems, Vol. 26 ,No. 5,DOI: 10.1023/A:1016496916732.

[10] Rencher , A. C., "**Methods of Multivariate Analysis**" John Wiley & sons, New York , USA,1995,ISBN 0-471-41889-7.

[11] Taha , H. A. ,1992,"**Operation Research**: **An Introduction**" , Macmillan publishing company , division of Macmillan Inc ,ISBN 0-13-188923-0 .

[12] Timm, N. H., 2002,**"Applied Multivariate Analysis** " New York,Springer, pp. 496-511,ISBN 0-387-95347-7..

[13] Wassef, M. R., and Rafla, N.,2020, "**HLS Implementation of Linear Discriminated Analysis Classifier**." *2020 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, ,DOI: 10.1109/ISCAS45731.2020.9181270.

[14] Zeidan, A.,2012, "**Discriminated Analysis in View of Statistical and Operations Research Techniques** " Faculty of commerce , Al- Azhar university , Girls branch.