SPRINGER OPTIMIZATION AND ITS APPLICATIONS 15

K. Kosmidou · M. Doumpos · C. Zopounidis

Country Risk Evaluation

Methods and Applications

Springer

COUNTRY RISK EVALUATION

Springer Optimization and Its Applications

VOLUME 15

Aims and Scope

Optimization has been expanding in all directions at an astonishing rate during the last few decades. New algorithmic and theoretical techniques have been developed, the diffusion into other disciplines has proceeded at a rapid pace, and our knowledge of all aspects of the field has grown even more profound. At the same time, one of the most striking trends in optimization is the constantly increasing emphasis on the interdisciplinary nature of the field. Optimization has been a basic tool in all areas of applied mathematics, engineering, medicine, economics and other sciences.

The series *Optimization and Its Applications* publishes undergraduate and graduate textbooks, monographs and state-of-the-art expository works that focus on algorithms for solving optimization problems and also study applications involving such problems. Some of the topics covered include nonlinear optimization (convex and nonconvex), network flow problems, stochastic optimization, optimal control, discrete optimization, multi-objective programming, description of software packages, approximation techniques and heuristic approaches.

COUNTRY RISK EVALUATION

Edited By

KYRIAKI KOSMIDOU
Technical University of Crete, Greece

MICHAEL DOUMPOS
Technical University of Crete, Greece

CONSTANTIN ZOPOUNIDIS
Technical University of Crete, Greece

Springer

Kyriaki Kosmidou
Department of Economics
Aristotle University of Thessaloniki
Thessaloniki, Greece
kosmid@econ.auth.gr

Constantin Zopounidis
Department of Production
Engineering and Management
Financial Engineering Laboratory
Technical University of Crete
University Campus
731 00 Chania
Crete, Greece
kostas@dpem.tuc.gr

Michael Doumpos
Department of Production
Engineering and Management
Technical University of Crete
University Campus
731 00 Chania
Crete, Greece
mdoumpos@dpem.tuc.gr

ISSN: 1931-6828
ISBN: 978-0-387-76679-9 e-ISBN: 978-0-387-76680-5
DOI: 10.1007/978-0-387-76680-5

Library of Congress Control Number: 2008922673

Printed on acid-free paper

springer.com

Learning really nourishes the mind
Empedokli (Dil 17, 14)

Preface

Country risk analysis has been one of the major research topics within the fields of economics and finance during the past years. Specific consideration is given to the examination of the factors that are related to the economic and financial difficulties that countries face, as well as on the investment environments. According to Cosset et al. (1992), country risk is defined as the probability that a country will fail to generate enough foreign exchange in order to pay its obligation toward foreign creditors. Based on this definition, the importance of studying the country risk assessment, as well as the establishment of country risk analysis systems, becomes obvious.

Researchers have employed several quantitative analysis methodologies to develop appropriate country risk assessment models. All these studies intend to identify the relationship between country risk indicators and the level of country risk of the countries. They point out the challenges that need to be met and the future research directions that can be explored. Moreover, new issues in country risk assessment have arisen, mainly after the recent crises in Asia and South America.

This book reviews the existing research in country risk analysis and presents several modeling methods for developing country risk assessment models, including statistical and non-parametric methods. Special emphasis is given to the use of multicriteria decision aid methods (MCDA) that may be employed in the country risk assessment problem. Because classic statistical and other methods applied in the past were not always able to respond effectively and sufficiently to the country risk problem, researchers proposed multicriteria methods. MCDA methods are quite interesting and attractive alternatives because they provide an environment for the employment of a sufficient number of quantitative and qualitative factors for the country risk assessment.

The book is organized in four chapters as follows:

Chapter 1 provides a review of country risk definitions, as well as an overview of the most recent operational tools in country risk assessment.

Chapter 2 makes a presentation of MCDA classification methods, statistical and econometric classification methods, and non-parametric techniques. It describes the general characteristics of the methods as well as the model development process for each of them.

Chapter 3 presents several real-world applications of the methodologies described in Chapter 2 on the country risk assessment problem and analyzes their results.

Finally, Chapter 4 concludes the book and proposes future research directions for the country risk assessment problem.

Chania, Greece
February 2008

Kyriaki Kosmidou
Michael Doumpos
Constantin Zopounidis

Contents

Chapter 1
Introduction

The oil crises of the 1970s and the worldwide economic turmoil were the first post-war events that highlighted the importance of a global risk factor for the management of firms and organizations, as well as for the sustainable socio-economic development of countries. Despite the stabilization of the global economic and business environment during the 1980s and the development of the 1990s, this risk factor has not lost its importance. The rapid change toward a globalized environment has already highlighted, in several situations, the multiplicative effect that a socioeconomic turmoil at a national or regional level may have at the global level. The recent crises in Southeast Asia and South America are typical and clear examples of this finding.

Country risk analysis has evolved as a major research topic within the fields of economics and finance during the past three decades, focusing on the investigation of the economic and financial difficulties that countries face, the factors that are related to these difficulties and their impact on economic policymaking, as well as on the business and investment environments. The major significance of country risk analysis is clearly understood by the plethora of existing risk rating agencies that provide assessments of country risk (Erb et al., 1996).

Country risk has many facets, because the economic and financial development of a country and the difficulties that it faces may have different origins. In a general sense, country risk is often defined as the probability that a country will fail to generate enough foreign exchange in order to pay its obligation toward the foreign creditors (Cosset et al., 1992). This is a purely economic definition of country risk. Several researchers, however, have emphasized the necessity of defining country risk in a broader context that better represents the multidimensional character of country risk. In that regard, Mondt and Despontin (1986) argue that the economic dimension of country risk only shows the capacity of a country to service its debt, but its willingness to service its debt should also be considered in the analysis through the investigation of the political environment in the country, which defines its political risk. Calverley (1990) employed this context to define country risk as the potential economic and financial losses due to the difficulties that are raised from the macroeconomic and/or political environment of a country.

K. Kosmidou et al. (eds.) *Country Risk Evaluation,* DOI: 10.1007/978-0-387-76680-5_1,

Obviously, these definitions relate the concept of country risk to the obligations that a country faces toward its foreign creditors (international banks, organizations, other countries). Other researchers introduced in the analysis the investment perspective, i.e., the impact of a country's economic and sociopolitical environment to the decisions made by international firms to undertake significant investment projects in this country. In this context, Herring (1983), Kobrin (1986), and Ting (1988) referred to the macro (sociopolitical) risks and the micro risks that the international investors are facing. The macro risk arises from dramatic events such as wars, sectarian conflicts, revolutions, etc., as well as less dramatic events such as the country-wide imposition of price controls, tax increases or surcharges, etc. The micro risks concern circumstances involving industry, firm or project-specific cancellation of import and export licenses, discriminatory taxes, etc.

The above definitions clearly indicate the multiple facets of country risk, involving both the factors that are relevant (economic/financial, social and political factors) and those that are interested in country risk analysis (bank managers, firms, international organizations, policymakers).

The first attempts to establish country risk analysis systems have been made mainly by banking institutions. These attempts involved simply devising checklist systems based mainly on economic variables (Saini and Bates, 1978). However, this approach has been proven to be insufficient mainly due to its inability to establish a sound methodological framework for the selection and weighting of the variables (Burton and Inoue, 1983).

1.1 Statistical Approaches in the Assessment of Country Risk

The empirical literature on country risk assessment has developed separate bodies. This section presents a review of statistical methods applied to country risk assessment such as discriminant analysis, factor analysis, regression analysis, regression trees, cluster analysis, logit analysis, and principal component analysis.

Moreover, it briefly describes studies that illustrate (1) the relationship between debt servicing capacity and economic/political indicators, (2) the use of creditworthiness indicators in country risk ratings, and (3) alternative views of assessing country risk including expert judgment on political riskiness, bank lending policies, and ways to include country risk in the appraisal of international investments. Table 1.1 provides a list of commonly used indicators for country risk analysis, whereas Table 1.2 summarizes several studies using statistical methods for country risk analysis. Figure 1.1 highlights the most important country risk indicators.

1.1.1 Debt Reschedulings

Since the second oil price shock of 1979–1980, debt servicing problems of countries with large external debt have increased the interest of banks and other international

Table 1.1 List of commonly used country risk indicators

Amortization rate	AR	Frequency of a governmental changes	FGC
Amortization/Debt	A/D	GDP growth	GDPG
Budgetary deficits	BD	GNP per capita	GNP/C
Capital inflows/Debt service payments	CI/DSP	Gross fixed capital formation/GDP	GFCF/GDP
Current account balance	CAB	Growth of GNP per capita	GNPCG
Current account balance/Exports	CAB/E	Imports/GDP	I/GDP
Current account balance/GDP	CAB/GDP	Imports/GNP	I/GNP
Debt in default	DID	Imports/Reserves	I/R
Debt service payments/Debt disbursement	DSP/DD	Infant mortality rate	IMR
Debt service payments/External debt	DSP/ED	Inflation growth	INFG
Debt service payments/Imports	DSP/I	Inflation	INF
Debt service payments/Reserves	DSP/R	Interest rates	IR
Debt service ratio	DSR	International liquidity	IL
Debt servicing capacity	DSC	Level of democracy	LD
Default history	DH	Loan commitments per capita	LCPC
Disbursed ext. debt/Debt service payments	DED/DSP	M2/Reserves	M2R
Disbursed ext. debt/Exports	DED/E	Money supply growth rate	MSGR
Domestic credit/GDP	DC/GDP	Political instability	PI
Economic development	DEV	Propensity to invest	PTI
Exchange rate differential	ERD	Rate of increase in consumer prices	ICP
Exports	E	Real exchange rate	RER
Export growth rate	EGR	Reserve position in IMF/Imports	RIMF/I
External debt/Exports	ED/E	Reserves growth rate	RGR
External debt/GDP	ED/GDP	Short-term debt/Exports	STD/E
External debt/GNP	ED/GNP	Short-term debt/External debt	STD/ED
Foreign aid per capita	FA/C	Sovereign credit rating	SCR
Foreign domestic invest. per capita outstanding	FDI/C	Stock returns	SR

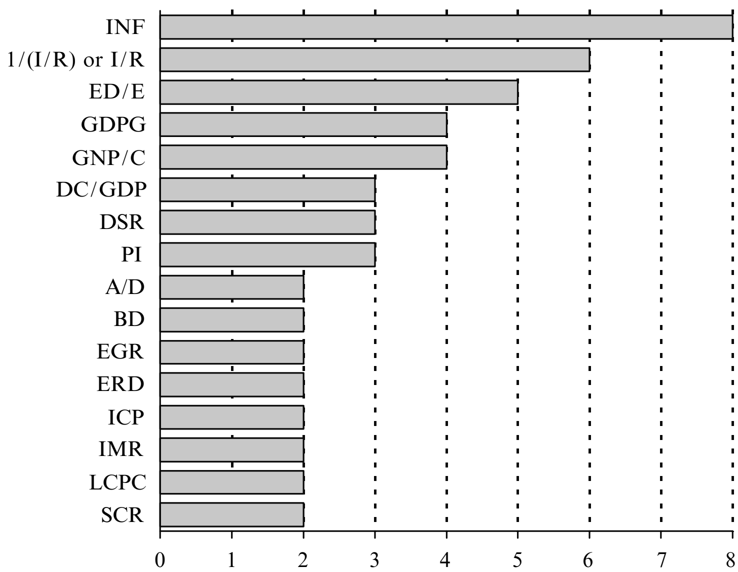

Fig. 1.1 Country risk indicators most often found significant in country risk studies (frequencies)

Table 1.2 Summary of studies based on statistical techniques for country risk analysis

Studies	Period	Methods	Dependent	Important
Frank & Cline (1971)	1960–1968	DA	DR	DSR, I/R, A/D
Dhonte (1975)	1959–1971	PCA	DR	DSP/DD, DSP/ED
Grinols (1976)	1961–1974	DA	DSD	DSP/R, DED/DSP, DSP/I, ED/GDP, ED/E
Feder & Just (1977)	1965–1972	LA	DR	DSR, I/R, A/D, GNP/C, CI/DSP, EGR
Sargen (1977)	1960–1975	DA	DR	INF, DSP/E
Mayo & Barrett (1978)	1960–1975	LA	DSD	DED/E, 1/(I/R), GFCF/GDP, I/GDP, RIMF/I, ICP
Saini & Bates (1978)	1960–1977	DA, LA	DR	ICP, MSGR, CAB/E, RGR
Abassi & Taffler (1982)	1965–1972	PCA, DA	DR	LCPC, ED/E, INF, DC/GDP
Taffler & Abassi (1984)	1967–1978	DA	DR	LCPC, ED/E, INF, DC/GDP
Feder & Uy (1985)	1979–1983	RA	SCR	GDPG, EGR
Burton & Inoue (1987)	1968–1977	DA	Foreign asset expropriation	FDI/C, FA/C, PI, BD, GDPG, INF, GNP/C
Citron & Nickelsburg (1987)	1960–1983	LA	DR	PI, IL
Cooper (1987)	1983	CA, DA	DR	GDPG, INF, ED/E, STD/ED, STD/E, 1/(I/R)
Mumpower et al. (1987)	1983–1985	FA RA	PRR	ERD, INF, IMR
Cosset & Roy (1988)	1983–1985	RT	PRR	ERD, INF, IMR
Brewer & Rivoli (1990)	1987		SCR	FGC
Cosset & Roy (1991)	1987	RA, RT	SCR	GNP/C, PTI
Balkan (1992)	1971–1984	PA	DR	LD, PI, DSR, I/GNP, AR, ED/GNP, 1/(I/R), GNPCG
Lee (1993)	1986	RA	SCR	ED/GNP, GDPG, DC/GDP
Cantor & Parker (1996)	1987–1994	RA	SCR, bonds yields	GNP/C, INF, ED/E, DEV, DH
Haque et al. (1996)	1980–1993	RA	SCR	1/(I/R), CAB/GDP, INFG
Ramcharran (1999)	1992–1994	RA	Secondary market prices of loans	SCR, DSC, DID
Kaminsky & Schmukler (2002)	1999–2000	RA	Stock returns, bonds yields	SCR, IR
Reinhart (2002)	1970–1999	PA	Currency crises, defaults	RER, SR, E, M2R, CAB, BD

Notes: DA = Discriminant analysis, LA = Logit analysis, PA = Probit analysis, RA = Regression analysis, PCA = Principal components analysis, FA = Factor analysis, RT = Regression trees, DR = Debt rescheduling, PRR = Political risk ratings, DSD = Debt servicing difficulties.

institutions, governments and the public. This section surveys the studies focused on statistical techniques for identifying the determinants of a country's debt servicing capability. More specifically, the focus is on (1) the factors that affect a country's ability and willingness to repay its external debt and describe its debt-servicing difficulties, (2) the prediction of debt reschedulings, and (3) the analysis of the relationship between economic factors and debt servicing capacity and the impact of economic determinants in developing country creditworthiness indicators.

Frank and Cline (1971) used discriminant analysis to investigate the ability of eight indicators to identify debt servicing difficulties. The variables selected were the debt service ratio, export fluctuations, compressibility of imports, the imports to GNP ratio, the imports to reserves ratio, the amortization to debt ratio, GNP per capita, and the growth of exports. Covering the period 1960–1968, their data sample contained 145 observations on 26 countries, of which 13 were rescheduling cases in eight countries. Frank and Cline tested their indicators on a binary-valued dependent variable consisting of rescheduling and non-rescheduling cases. Three variables were found significant, including the debt service ratio, the imports to reserves ratio, and the amortization to debt ratio that generated fairly low error rates and explained correctly 10 of the 13 rescheduling cases. Furthermore, they found that with discriminant analysis, it is possible to obtain a very high prediction rate using only two factors: the debt service ratio and the average maturity of debt.

Dhonte (1975) used principal components analysis to analyze debt rescheduling. He selected 13 cases of debt rescheduling between 1959 and 1971 and compared them with a sample of 69 non-rescheduling countries in 1969. He found four indicators to be the most significant for the first principal component. In addition, for the second principal component, Dhonte found two more significant indicators, namely the debt service payment/debt disbursement ratio and the debt service payments/external debt ratio. Dhonte concluded that in order to avoid a debt-servicing problem, a balance must be maintained between a debtor's involvement in debt and the terms on which a debt is accumulated.

Grinols (1976) applied discriminant and discrete analysis to a broader set of variables and to an expanded sample of countries, with observations covering the 1961–1974 period, in order to describe the difficulties in servicing external debt. This study identified five statistically significant variables: the debt service payments to reserves ratio, the disbursed external debt to debt service payment ratio, the debt service payments to imports ratio, the external debt to the GDP ratio, and the external debt to exports ratio. The estimated discriminant function showed error rates within the test sample, which were almost 50% lower than the results obtained by Frank and Cline (1971).

Feder and Just (1977) used logit analysis in order to reinvestigate the significance of the indicators used by Frank and Cline, Grinols, and Dhonte. Their sample referred to 238 observations on 30 countries (21 rescheduling cases involving 11 countries) for the period 1965–1972. From the eight indicators examined, six variables were found to be significant: the debt service ratio, the import to reserves ratio, the amortization to debt ratio, the per capita income, the capital inflows to debt service payments ratio, and the real export growth rate. The estimated logit model achieved lower error rate compared with the results obtained in the previous studies.

Mayo and Barrett (1978) designed an early warning model for the U.S. Export-Import Bank (Eximbank) based on earlier studies. Their analysis involved the development of a logit model considering a sample of 48 countries during the period 1960–1975. In contrast with the previous studies that focused on debt rescheduling, Mayo and Barret considered debt servicing difficulties in a broader sense. The objective of their model was to predict debt servicing difficulties for up to five years prior to the time where debt difficulties would be evident. Among the indicators found to be statistically significant, the six that provided the best predictive results were the disbursed external debt to exports ratio, the reserves to imports ratio, the gross fixed capital formation to GDP ratio, the imports to GDP ratio, the reserve position in the International Monetary Fund to imports ratio, and the rate of increase in consumer prices.

Sargen (1977) used two conceptual approaches to analyze past debt reschedulings. The first approach was similar to the approaches employed in the other studies and assumed that reschedulings arise from fluctuations in export earnings, which lead to a rapid accumulation of external debt. The second approach treats reschedulings as a monetary phenomenon. In this approach, inflation and an overvalued exchange rate were assumed to increase the demand for imports and to cause export stagnation, which in turn leads to a rapid build-up of external debt. To test the validity of each approach, Sargen applied discriminant analysis to six indicators, covering the period 1960–1975, to differentiate rescheduling from non-rescheduling cases. He found the inflation rate and the ratio of scheduled debt service payments to exports to be the two most significant explanatory indicators.

Saini and Bates (1978) began with the presumption that there was no a priori reason for expecting the logit analysis to be superior to the discriminant analysis. Their sample contained data on 25 countries covering the time period 1960–1977, from which 13 countries either faced rescheduling and/or secured balance of payments support loans in the given period. They utilized a modified dependent variable that included reschedulings and balance of payments support loans. The indicators used in the analysis were selected according to their statistical significance in at least one of the previous studies, with the exception of those indicators that contained debt data. Several proxy variables for external indebtedness were constructed. Their results indicate that the four most significant explanatory variables include the growth rate of consumer prices, the money supply growth rate, the ratio of the adjusted cumulative current account balance to exports, and the growth rate of reserves.

Abassi and Taffler (1982) used discriminant analysis in order to evaluate country risk. Their sample of 1140 observations on 95 countries for the period 1965–1972 contained 55 rescheduling cases referring to 14 countries. Their dependent variable did not exclude voluntary debt reschedulings, nor did it include balance of payments support and/or bridge loans. They considered 42 indicators in their analysis, regarding the foreign exchange sector, country debt, and the domestic economic situation. They incorporated several novel features into their model. First, to identify the degree of intercorrelation among variables, they used principal components analysis. Second, to correct for serial correlation and to obtain an unbiased estimate of the true classification error rate, the model was calculated using a step-wise approach for variable selection. The model was estimated for the 1967–1977 period and tested on the 1978 data. The final variable set consisted of the following most important indicators: new loan commitments per capita, external debt to exports ratio, rate of inflation, and domestic credit to GDP ratio.

Taffler and Abassi (1984) developed a discriminant model in order to predict debt reschedulings among developing countries. They differentiated their previous work by incorporating into their model both monetary policy and debt servicing capacity indicators. The 42 variables used in the analysis were the ones that had been found useful in other studies. Economic information referring to 95 developing countries was gathered for a 12-year period up to 1978. Data from 1967 to 1977 were used to fit the model, which was then tested using the data of 1978, in order to distinguish possible rescheduling countries in 1979. Their model consisted of the following four most important variables: commitments per capita, debt to exports, average rate of inflation, and domestic credit to gross domestic product. Although their model appeared relevant to the least developed countries, some difficulties were faced with regard to the countries experiencing short-term problems.

Cooper (1987) utilized cluster analysis and multiple discriminant analysis to distinguish countries that were likely to seek a rescheduling of their debt. Cluster analysis was used to partition countries into two groups. The first group comprised countries that did not seek any rescheduling of their international debt obligations in 1983, and the second group comprised countries that rescheduled all or part of their debt in 1983. In total, eight explanatory variables were considered in the analysis, including average GDP growth, inflation, the external debt ratio, the ratio of short-term debt to total external debt, the ratio of short-term debt to exports, the ratio of reserves to imports, and two debt-service ratios. The results indicated that the cluster analysis was quite efficient in providing accurate predictions. The eight explanatory variables were also used in developing a discriminant analysis model. The main conclusion was that, similar to cluster analysis, discriminant analysis also performed well in its predictions.

Haque et al. (1996) provided an empirical analysis of the economic determinants of developing country creditworthiness models for more than 60 developing countries during the period 1980–1993. Their study extended the earlier analyses by examining the behavior of three creditworthiness series over the longest time period used to date. The data set consisted of the credit ratings constructed by Institutional Investor, Euromoney, and the Economist Intelligence Unit. Two different theoretical approaches were used to model country default risk. The debt-service capacity approach regarded default as arising out of an unintended deterioration in the borrowing country's capacity to service its debt. On the contrary, the cost-benefit approach viewed the rescheduling (or default) of a country's external debt as a rational choice based on an assessment of the costs and benefits of rescheduling of repudiation. Their empirical results indicated that economic indicators played a key role in determining a developing country's credit rating. The most important domestic economic variables influencing country credit ratings were found to be the ratio of nongold foreign exchange reserves to imports, the ratio of the current account balance to GDP, the country's rate of growth, and inflation. All country ratings were adversely affected by increases in international interest rates, independent of the domestic economic fundamentals. A country's regional location and the structure of its exports were also found to be important.

1.1.2 Political Factors

A country's decision to reschedule its external debt reflects not only its economic circumstances, i.e., its ability to meet its obligations, but also its willingness to service these obligations. The latter reflects the political environment of the debtor country in that the decision to reschedule is a political decision. Also, many leading international institutions analyze and publish country risk and creditworthiness ratings based on the political factor impacts. A review of the past studies that incorporate not only the economic factors but also political ones is presented in this section.

Feder and Uy (1985) attempted to explain cross-sectional and inter-temporal variation in credit ratings based on Institutional Investor data. The study consisted of two main models. First, a regression analysis model was developed to determine the significant explanatory variables. The results showed that all variables were statistically significant. The authors also examined changes over time in the impact of economic indicators on creditworthiness and found that there was a significant difference between the periods 1979–1981 and 1982–1983. The second model simulated the evolution over time of a hypothetical economy by generating the time profile of exports, imports, reserves, GNP, external debt, and consumption. The obtained results suggested that a higher rate of growth of GDP, holding export growth constant, improved the initial creditworthiness rating, and an increase in the rate of growth of exports significantly strengthened creditworthiness.

Citron and Nickelsburg (1987) proposed a logit model of country risk that incorporated not only economic but political variables, too. They incorporated into their model the political instability indicator, which was proxied by the number of changes of government over a five-year period. More specifically, they modeled in a very simple way the factors that increase the change of default. They noticed that when a government was characterized by an unstable environment, for example, by the exact time after a new government takes over, the increase in government welfare through spending depends essentially on domestic purchases. Even though payment of debts was beneficial, if the unstable government wanted to retain its power, it had to make sure first that its expenditures were directed toward those who might overthrow it. So, if the debtor country has the opportunity to tax the foreign creditor through rescheduling or other reduction in debt payments, this might be positive for the government's welfare. Their model was estimated simultaneously for five countries (Argentina, Brazil, Mexico, Spain, and Sweden) from 1960 to 1983. Their empirical results showed that political instability was a very important variable that has to be taken into account in country risk analysis because it affected a government's willingness to service its debt payments. Furthermore, international liquidity was found to be highly significant.

Mumpower et al. (1987) studied the professional analysts' judgments of the political riskiness using factor and regression analysis. The dependent variable was the political risk ratings of 49 countries as given by the annual survey of the Association of Political Risk Analysis for the years 1983–1985. Factor analysis identified three significant explanatory variables: the exchange rate differential, the estimated inflation rate, and the infant mortality rate. Furthermore Mumpower et al. (1987) broke down their sample by geographic area (safe and unsafe) in order to test the stability of the model. Finally, they made a comparison between the ratings of the experts and the ratings of undergraduate students of political science and found that they were closely parallel. Thus, they concluded that even though experts have an advantage regarding the quantity of information, their estimations are almost equivalent to those made by naïve subjects and by a simple linear model involving only few variables.

Burton and Inoue (1987) developed a country risk appraisal model of foreign asset expropriation in developing countries. They used a multiple discriminant analysis in their model in order to clarify the important economic, political, and environmental variables that differentiate the countries that expropriated or not the assets of foreign firms over the period 1968–1977. The dependent variable referred to whether an expropriation event occurred. The most significant variables found are the foreign domestic investment per capita outstanding, foreign aid per capita, political instability, budgetary deficits, GDP growth rate, inflation, and GNP per capita. The accuracy of the model was improved by the incorporation of dummy variables regarding income level groups and regional groups.

Cosset and Roy (1988) used the regression tree technique on the same data set used by Mumpower et al. (1987) in order to study the experts' judgments of political riskiness. The assessment of this technique led them to the selection of the same statistically significant variables identified by Mumpower et al. (1987) regarding the estimated inflation rate, the infant mortality rate, and the exchange rate differential. The infant mortality rate was used twice at different stages of the regression tree. Through this regression tree a significant improvement in the correlation coefficient was achieved compared with the results of Mumpower et al. (1987). This improvement shows that a regression tree extracts more information from the data set, thus avoiding the problem of multicollinearity. Furthermore, regression trees do not allow variable interaction and do not impose any assumptions on the distribution of the prediction variables.

Cosset and Roy (1991) replicated in their study both Euromoney and Institutional Investor's country risk ratings (scores of 71 countries as reported in the September 1987 issues). The explanatory variables used were the ones derived from earlier studies and theoretical models upon the international borrowing in the presence of default risk. They applied two statistical techniques in their study: linear regression analysis and regression trees. It was found that the level of per capita income and propensity to invest are the two variables that affect positively the rating of a country. The results also reveal that the ability of the models to reproduce the two country risk measures is very similar.

Brewer and Rivoli (1990) in order to determine creditworthiness focused on the effect of political instability as well as on the impact of some economic variables. Country creditworthiness data were taken from the 1987 Institutional Investor and Euromoney ratings. The explanatory variables included several measures of political instability and armed conflict, but only two economic measures: the ratios of current account to GNP and external debt to GNP. Cross-sectional analysis was performed using data for the 30 most heavily indebted developing countries. The data on economic variables were for the year 1987, and explanatory variables were computed over the 1967–1986 period. The results showed that, whereas the frequency of a change in government regime was significant as a proxy for political stability, two other variables, proxying the degree of armed conflict and political legitimacy, were not significant.

Balkan (1992) incorporated into his empirical work two dimensions of the borrower's political environment: the level of democracy and the level of political instability. For each of these two dimensions, an index was created using the Banks Cross-National Time Series Data Archive (1986). These indexes were included in a probit model, along with other economic variables commonly used in previous empirical works. Annual data were used for 33 countries over the period 1971–1984. The dependent variable referred to the situation where a country had rescheduled or not its sovereign external debts in a given year as reported by the World Bank. According to the results, an inverse relationship between rescheduling probabilities for a given country and its level of democracy and a direct relationship between the rescheduling probabilities and the political instability level were found. Furthermore, most of the economic variables were found to be statistically significant. In particular, the probability of rescheduling was found positively related to debt service, the ratio of imports to GNP, and the ratio of debt to GNP, whereas it was negatively related to the ratio of international reserves to imports, the amortization rate, and the growth of GNP per capita.

Lee (1993) examined the effects of both economic and political variables in explaining country risk ratings. His sample consisted of 29 heavily indebted countries. Institutional Investor and Euromoney were again the main sources that provided the sample data. The explanatory variables included three economic variables: the ratio of external debt to GNP, per capita GDP growth, and the ratio of domestic public debt to GDP. Other debt-service variables such as the ratios of total debt to exports and reserves to imports were also taken into consideration. The results suggested that creditworthiness indicators were explained mainly by the countries' economic performance, rather than by their political situation.

1.1.3 Alternative Views in the Assessment of Country Risk

The lending policies and the ways in which country risk is assessed by bankers, the determinants of supply and demand for sovereign loans, the impact of international lending and borrowing in the world income redistribution, and the answer to how country risk may be included in the evaluation of investment projects present some of the alternative views in the evaluation of country risk.

Agmon and Deitrich (1983) studied an alternative approach of country risk, referring to the international lending and income redistribution. Traditional approaches upon creditworthiness of the borrowing countries had no impact on credit granting decisions by international banks. Agmon and Deitrich in their effort to go beyond these traditional approaches presented a model in order to explain international borrowing and lending activities. The basic assumption of this model involved the use of activities like lending and borrowing as an appropriate way of influencing the redistribution of income in the world. It was found that loan servicing was dependent on the competence of the borrowing country to tax the lending country indirectly through financial intermediaries. Furthermore, factors affecting risks to loans are those that determine the responsibility of the government of the lending countries to subsidize wealth transfers to borrowing countries.

Heffernan (1985) studied country risk analysis from the demand and supply of sovereign loans points of view. With the term sovereign loans, Hefferman refers to the loans made directly either to the public sector or to a private debtor having the guarantee that they will be paid by the public entity. He developed a general model of the demand and supply of loans made to the public sector of a third-world country. On the supply side, emphasis was given to the need of lenders to taking into account the sovereign loans as part of an optimal investment decision. On the demand side, a life-cycle hypothesis of developing economies was made in order to identify the dependence of the demand for external debt on variables such as domestic rate of savings, capital rental rate and its relation to the world interest rate, and the value of the country's domestic output. Endogenizing the probability of default expanded the model.

Shapiro (1985) studied the conditions under which banks are subject to currency and country risks. Banks face currency risks when they lend to foreign firms and governments. They are trying to overpass these risks either by denominating and funding their loans in the foreign currency or by denominating their foreign loans in U.S. dollars. Following this practice, banks are protected, shifting any risk associated with exchange rate fluctuations to the borrowers. Currency risk is converted into credit risk to the extent that changes in currency values can affect the ability or willingness of foreign borrowers to repay their loans. In the case where the government is the borrower, credit risk becomes country risk. Shapiro concluded that currency risk is dependent on the rate of both domestic and foreign inflation, the deviations of purchasing power parity, and on the effect of these deviations upon the firm's and the country's dollar-equivalent cash flows. On the other hand, the variability of the country's terms of trade and the government's willingness to permit the national economy to adjust rapidly to economic changes are important situations that determine country risk.

Bird (1986) examined the lending policies of international banks. The variables considered when assessing country risk from the banking point of view were subdivided into economic and political categories. Political indicators tended to be rather judgmental, although some banks employed political risk analysts and approached the question in a more objective and structured fashion. Whatever the exact mechanism used, most bankers felt that risk analysis is an "art" rather than a "science," and this has important implications in an environment where it is hoped to alter lenders' attitudes and perceptions. Bird suggested that benefits were to be gained from reforming the methods of country risk analysis used by the banks. The reforms were mostly addressed toward improving the underlying economic rationality and stability of bank lending decisions.

Somerville and Taffler (1995) attempted to explore the judgmental accuracy of bankers, in the context of country risk assessment, and they compared it with the performance of formal statistical models in terms of forecasting power over a one-year horizon. The Institutional Investor country credit ratings were taken as indicators of banker judgments. Creditworthiness was represented by a binary dependent variable and the training sample referred to the period 1980–1987. The results obtained using discriminant and logit analysis showed that the Institutional Investor credit ratings were biased toward an adverse view of the creditworthiness of less developed countries (LCDs) during 1987–1989. This “over-pessimism” contrasted with the findings of Taffler and Abassi (1984) of “over-optimism” at the beginning of the decade (i.e., 1980, 1983). In both periods, the average view of bankers appeared to have been biased and may possibly be interpreted as evidence of judgmental failings.

Nordal (2001) addressed the question of how country risk may be included in the valuation of investment projects. The main point of his research was that country risk indices might be modeled as stochastic processes and used as state variables when applying the contingent claims valuation methodology. The author specified a model that serves as a starting point when evaluating a variety of investment projects. Based on data covering a selection of country risk indices for oil-producing countries, he examined the properties of the model. He ended the research with a numerical example where the incentive for the investor to delay the investment decision was computed. This incentive was dependent on, among other elements, the current conditions in the country and the expected development, as well as on the volatility of a risk index.

1.1.4 Sovereign Credit Ratings

In recent years, the demand for sovereign credit ratings has increased dramatically. Several of the studies discussed in the preceding subsections considered sovereign credit ratings as an adequate measure of country risk and focused on the identification of the determinants of these ratings on the basis of economic indicators and political factors. However, in the light of the recent economic crises in Asia and South America, there has been some skepticism on the quality of these ratings, their impact on the financial markets, and ultimately on their ability to predict the crises. These questions constitute major research points in recent studies on country risk analysis.

Cantor and Packer (1996) presented the first systematic analysis focusing not only on the determinants of sovereign credit ratings but also on their impact to the bond markets. The authors focused on the ratings of Moody’s and Standard and Poor’s. The analysis was based on a sample of 49 countries rated by both Moody’s and Standard and Poor’s in their 1995 ratings. The countries were described by seven economic indicators and a dummy variable of their default history. At the first stage, a regression analysis was performed to investigate the relationship of the eight explanatory variables with the ratings of the two agencies. The results showed that the contribution of the explanatory variables in explaining the ratings is similar for both agencies. The most important variables included the per capita income, the inflation, the external debt, the economic development, and the default history. At the second stage, the analysis focused on the effect that the ratings have on bond spreads over U.S. Treasuries. The results showed that the ratings have considerable power in explaining bond yields, even though the financial markets are, generally, more pessimistic than the agencies for low-rated countries. Finally, using a sample regarding rating announcements for the period 1987–1994, they found that the changes in the ratings are followed by significant bond yield changes in the expected direction (downgrades in rating are followed by a rise of bond yields and vice versa).

Ramcharran (1999) conducted a similar analysis focusing on the case of the least developed countries (LDCs). The sample used in the analysis involved cross-sectional data on 27 LCDs for the period 1992–1994. The considered independent variables involved economic and political factors related to the country's ability and the willingness to service its debt and the Euromoney's sovereign credit rating. The dependent variable involved the price of each country's debt in the secondary market. The results of a regression analysis showed that sovereign credit rating is the most important determinant, followed by the debt-servicing capacity of LDCs and debt in default.

Kaminsky and Schmukler (2002) have also been involved with the sovereign ratings established by international agencies. Their analysis focused on the impact that these ratings have on financial markets. The motivation of this analysis was on the finding that often rating agencies tend to have a "procyclical" behavior (1999), i.e., they upgrade countries in good times and downgrade them in bad times, thus triggering market jitters contributing to the vulnerability that has been recently evident in the international stock markets. The data used in the analysis involved daily series of emerging markets bond index (EMBI) spreads, stock returns, interest rates, and credit ratings for 16 emerging markets over the period January 1999 to June 2000. Different regressions were performed to explore the changes in stock market prices and the bond spreads in terms of the changes in credit ratings and the changes in the U.S. interest rates. Similar to other studies (Cantor and Packer, 1996; Ramcharran, 1999; Reisen and Von Maltzan, 1999), the results show that rating changes significantly affect bond and stock markets. Furthermore, rating changes among emerging markets have a contagion effect triggering changes in bond spreads and stock returns in foreign countries, mainly within a specific geographical region. On the other hand, changes in the U.S. interest rates have more significant impacts for the countries whose ratings are low. Finally, it was observed that country rating upgrades take place after market rallies, whereas downgrades occur after market downturns, thus supporting the aforementioned finding on the "procyclical" behavior of rating agencies.

In contrast with the previous two studies that focused on the effects of rating on the financial markets, the study of Reinhart (2002) investigated the relationship among sovereign credit ratings, currency crises, and default. The author considered the credit ratings of Institutional Investor, Moody's, and Standard and Poor's during the period 1970–1999. The sample used in the analysis included 113 defaults and 151 currency crises, 135 of them in emerging market economies. The majority of the defaults (84%) in emerging economies in the sample has been associated with currency crises, but the converse was found not to be true. For developed economies, the author did not find evidence of any connection between currency crises and default. Using a probit model, the author found that the significant indicators for predicting currency crises included the real exchange rate, the stock returns, the exports, the M2/reserves ratio, the current account balance, and the overall budget deficit as a percent of GDP. A similar list of indicators was also found significant for predicting defaults. Furthermore, the results of the developed probit model showed that sovereign credit ratings systematically fail to predict currency crises but do considerably better in predicting defaults. Similar results have also been found by Goldstein et al. (2000) who examined the links among currency and banking crises and changes in sovereign credit ratings by Institutional Investor and Moody's for 20 countries. Reinhart (2002) attributed the inability of sovereign credit ratings to predict currency crises to the fact that currency crises are generally difficult to predict.

1.1.5 Important Issues in Statistical Country Risk Analysis

Obviously, the statistical approaches to country risk assessments have been widely used in the past and contributed positively in highlighting several important aspects of country risk analysis and the indicators that are involved. However, these approaches have several limitations in issues such as the specification of the dependent variable, data requirements and availability, model specification of the dependent variable, and forecasting ability. According to Saini and Bates (1984), the following five drawbacks are related to the statistical techniques and the related studies of the past:

- The definition of the dependent variable regarding the classification of the countries into rescheduling and non-rescheduling ones is not always a realistic approach. Countries have options other than formal reschedulings when they are facing debt-servicing problems. There are also substitutions for formal reschedulings such as debt refinancings and restructurings, etc. Furthermore, the definition of the dependent variable overlooks voluntary and nonvoluntary reschedulings. Voluntary reschedulings happen when balance of payments problems are not occurring.
- The reliance on debt information that is incomplete at least as far as it concerns the long-term case. For this reason, it is very important for researchers that introduce variables containing debt statistics into their studies to take into account that debt information on external debt is incomplete (Nowzard and William, 1981), mainly because of the lack of information on short-term debt. Furthermore, the reliance on debt information may decrease from the usefulness of empirical investigations.
- The statistical restrictions, such as: a) the reduction of the original data set to one with a smaller dimensionality that is useful only in the case where the interpretation of the newly constructed variables is meaningful, b) the determination of the importance of the explanatory variables and the problem of how to discard variables once included in the analysis, c) the difficulty in interpreting the obtained results mainly because of the lack of any explicit procedure for the selection of the value that distinguishes rescheduling from non-rescheduling cases, etc.
- Three weaknesses regarding model specification: a) the exclusion of important social and political factors that may lead to debt-servicing difficulties from the

analysis, b) the assumption of stable statistical relationships across countries regarding the basic structure and behavioral pattern for all countries, c) the overlooking of the dynamic nature of the world economy.
- The poor predictability of reschedulings of the statistical models, as statistically significant variables were found to be inadequate in making accurate predictions.

1.2 Multicriteria Analysis in the Assessment of Country Risk

To overcome the limitations and difficulties of the statistical approaches for country risk assessment, new methodological tools should be introduced in this field. Among them, multicriteria decision aid (MCDA) (Roy, 1996), an advanced field of operations research, seems to be well-suited to the analysis of country risk. MCDA is devoted to the development of appropriate quantitative methodologies for supporting the decision-making process in complex real-world problems that involve the consideration of a diversified and conflicting set of decision-making variables (evaluation criteria). This MCDA paradigm is well-suited to the multidimensional nature of country risk assessment. MCDA methods are free of statistical assumptions on the examined country risk data, they enable the incorporation of the decision makers' (managers of banks and international institutions, policymakers, investors) judgment policy into the analysis of country risk, they are capable of handling qualitative social and political factors, and they are easily updated taking into account the dynamic nature of the world economy, adapting to the changes in the decision environment. The MCDA methodologies already applied in country risk assessment studied the problem either from the ranking point of view, the portfolio construction point of view, or from the classification point of view. Several studies have proposed different approaches of multicriteria analysis in the assessment of country risk to support the decision-making process in problems related to country risk assessment. In particular, most of the studies on the use of MCDA methods in country risk analysis (including the studies discussed in this section) involve the development of risk assessment models. Risk assessment models often constitute a major part of integrated country selection models, i.e., models for selecting countries that are suitable to finance or to invest in. An example of an approach that deals with risk assessment in the context of a country selection process can be found in Kugel (1973).

Mondt and Despontin (1986) proposed a model to evaluate country risk in an interactive way using multiobjective linear programming. They applied the perturbation method (Vincke, 1976), a variant of the well-known STEM method (Benayoun et al., 1971), to determine the proportion of each country in an "optimal" loan portfolio of a bank. The aim of their analysis was to maximize the return of the portfolio and to minimize the corresponding risk. The method was applied to a sample of 10 countries and several risk dimensions such as the inflation risk, the exchange risk, the political risk, the social risk, and the growth risk were taken into consideration. The perturbation approach was employed as a scenario analysis approach to provide the decision maker with "what if" analysis results. Although this approach shows the contribution of each country to the risk of the whole portfolio, it does not provide an overall country risk rating according to the creditworthiness of the countries.

Tang and Espinal (1989) developed a multiattribute quantitative model to assess country risk, both on a short and medium-long term basis. The model developed in this research considered only quantifiable social and economic factors that could be available or estimated from reliable international sources. The model was applied to a sample of 30 developed and developing countries. The Delphi method (Lindstone and Turoff, 1975) was used in the selection of the relevant variables (external repayment capability, liquidity, per capita and population increases, purchasing power risk) and in the estimation of their relative weights (significance) in the analysis. The results showed that the most significant country risk indicator both for short and medium-long terms was the external repayment capability of a country. Comparing the results of the quantitative model with those obtained from two international banks, the validity of the model was evident regarding its broad consistency.

Oral et al. (1992) proposed a generalized logistic regression model to assess country risk. The parameters of the proposed model were estimated through a mathematical programming formulation that is able to consider the impacts of countries of different geographical regions or even countries with different political and economic characteristics. This model reproduced the country risk rating scores of Institutional Investor and it was applied to a sample of 70 countries for the years 1982 and 1987. The results obtained by the proposed generalized logistic model were compared with those obtained by two statistical models, namely logistic regression and regression trees. A comparison of the three methods indicated the superiority of the new method over the statistical models, with respect to both the estimation and validation samples. Regarding the importance of country risk indicators, the three models provided similar results. The generalized logit model indicated that for both years 1982 and 1987, the most important indicators were the net foreign debt/exports, the GNP per capita, and the investment/GNP. Furthermore, it was found that developed countries and countries geographically located in Southeast Asia were those that experienced low risk, whereas countries located in Central America experienced high risk.

Cosset et al. (1992) applied a preference disaggregation methodology for the evaluation of country risk, based on the MINORA decision support system (Multicriteria INteractive Ordinal Regression Analysis; Siskos and Yannacopoulos, 1985), which incorporates the multicriteria method UTASTAR, a variant of the UTA method (UTilités Additives; Jacquet-Lagrèze and Siskos, 1982). The MINORA system was applied to a sample of 76 countries for the year 1986, in order to develop a ranking model of the countries according to their ability to service their foreign currency loans. Using a sample of 22 reference countries, an additive utility model was interactively developed, which consistently represented the preferences of a decision maker. The results obtained from the proposed model showed that European countries, the United States, Canada, and Japan had the best performance regarding their creditworthiness, whereas countries such as Nigeria, Argentina, etc., were the most risky ones. Furthermore, the most important determinants of sovereign creditworthiness were found to be the GNP per capita, the propensity to invest, the current account balance on GNP. In conclusion, this process appears to be well adapted to deriving and updating a country risk evaluation model.

Cook and Hebner (1993) presented a multicriteria approach to country risk evaluation based on the pure ordinal model developed by Cook and Kress. The data used in the analysis were obtained from the Japan Bond Research Institute for 100 countries evaluated over 14 criteria. The results showed that investors with heterogeneous projects obtain different country risk rankings from one another (because they possess heterogeneous criteria importance vectors). Furthermore, the country risk rankings obtained using the multicriteria approach differed from those obtained by a quantitative fixed weighting approach. The ability of this study's multicriteria approach to determine the set of risk criteria weights optimally rather than relying upon exogenously given weights allows investors to formulate individualized country risks ratings, which reflect the unique risk sensitivities possessed by each of their projects.

The assessment of country risk was applied to the data of Tang and Espinal (1989). More specifically, in their study they applied three multicriteria methods, namely the UTASTAR method for developing a ranking country risk model, the UTADIS method (UTilités Additives DISscriminantes; Jacquet-Lagrèze, 1995), and a variant of the UTADIS method (UTADIS I; Zopounidis and Doumpos, 1998) for developing country risk classification models. The most important indicators for all the three methods were found to be the per capita income and population increases, the current account imbalance as a percentage of gross external revenues (GER) during recent periods, the imbalance between external debit and credit interest as a percentage of GER during recent years, the current account imbalance as a percentage of the GER increase during the recent period, and the gross international reserves as a percentage of gross external expenditures. The results obtained from both the ranking and the classification approaches of country risk analysis were very satisfactory as they were highly consistent with the estimations of the two lending institutions considered in the analysis of Tang and Espinal (1989).

Chapter 2
Review of Methodologies

The review of the previous chapter shows that country risk analysis is often based on the development of models to discriminate between high-risk countries (e.g., rescheduling) and low-risk ones.

This chapter is focused on the methods that can be used to develop such models. Special emphasis is given to non-parametric techniques from the field of MCDA. The considered methods include the UTADIS method and the MHDIS method (Multi-group Hierarchical DIScrimination). Both methods lead to the development of additive models that can be used to classify a set of alternatives (e.g., countries) into q predefined ordinal groups:

$$C_1 \succ C_2 \succ \cdots \succ C_q$$

where C_1 denotes the group consisting of the most preferred alternatives and C_q denotes the group of the least preferred alternatives. Within the country risk context, C_1 consists of the low-risk countries, whereas C_q consists of the high-risk ones.

The subsequent sections of this chapter discuss in detail all the model development aspects of the two methods as well as all the important issues of the model development and implementation process.

In addition, the two MCDA methods other techniques also discussed, including statistical methods, neural networks, rule induction and decision trees, fuzzy sets, and rough sets.

2.1 The UTADIS Method

2.1.1 Criteria Aggregation Model

The UTADIS method was first presented by Devaud et al. (1980), and some aspects of the method can also be found in Jacquet-Lagrèze and Siskos (1982). Jacquet-Lagrèze (1995) used the method to evaluate R & D projects, and during the past

K. Kosmidou et al. (eds.) *Country Risk Evaluation,* DOI: 10.1007/978-0-387-76680-5_2,

few years the method has been widely used for developing classification models in financial decision making problems (Zopounidis and Doumpos, 1998, 1999a, b; Doumpos and Zopounidis, 1998; Zopounidis et al., 1999). Recently, the method has been implemented in multicriteria decision support systems, such as the FINCLAS system (Zopounidis and Doumpos, 1998) and the PREFDIS system (Zopounidis and Doumpos, 2000a).

The UTADIS method is a variant of the well-known UTA method (UTilités Additives). The latter is an ordinal regression method proposed by Jacquet-Lagrèze and Siskos (1982) for developing decision models that can be used to rank a set of alternatives from the best to the worst ones.

Within the sorting framework described in the introductory section of this chapter, the objective of the UTADIS method is to develop a criteria aggregation model used to determine the classification of the alternatives. Essentially this aggregation model constitutes an index representing the overall performance of each alternative along all criteria. The objective of the model development process is to specify this model so that the alternatives of group C_1 receive the highest scores, while the scores of the alternatives belonging to other groups gradually decrease as we move toward the worst group C_q.

Formally, the criteria aggregation model is expressed as an additive utility function:

$$U(\mathbf{g}) = \sum_{i=1}^{n} p_i u_i(g_i) \tag{2.1}$$

where:

$\mathbf{g} = (g_1, g_2, \ldots, g_n)$ is the vector of the evaluation criteria.

p_i is a positive scaling constant indicating the significance of criterion g_i $(p_1 + p_2 + \cdots + p_n = 1)$.

$u_i(g_i)$ is the marginal utility function of criterion g_i.

The marginal utility functions are monotone functions (linear or nonlinear) defined on the criteria's scale, such that the following two conditions are met:

$$\left.\begin{aligned} u_i(g_{i*}) &= 0 \\ u_i(g_i^*) &= 1 \end{aligned}\right\}$$

where g_{i*} and g_i^* denote the least and the most preferred value of criterion g_i, respectively. These values are specified according to the set of the alternatives under consideration, as follows:

- For increasing preference criteria (criteria for which higher values indicate higher preference, e.g., return/profitability criteria):

$$g_{i*} = \min_{\forall \mathbf{x}_j \in A} \{g_{ji}\} \quad \text{and} \quad g_i^* = \max_{\forall \mathbf{x}_j \in A} \{g_{ji}\}$$

- For decreasing preference criteria (criteria for which higher values indicate lower preference, e.g., risk/cost criteria):

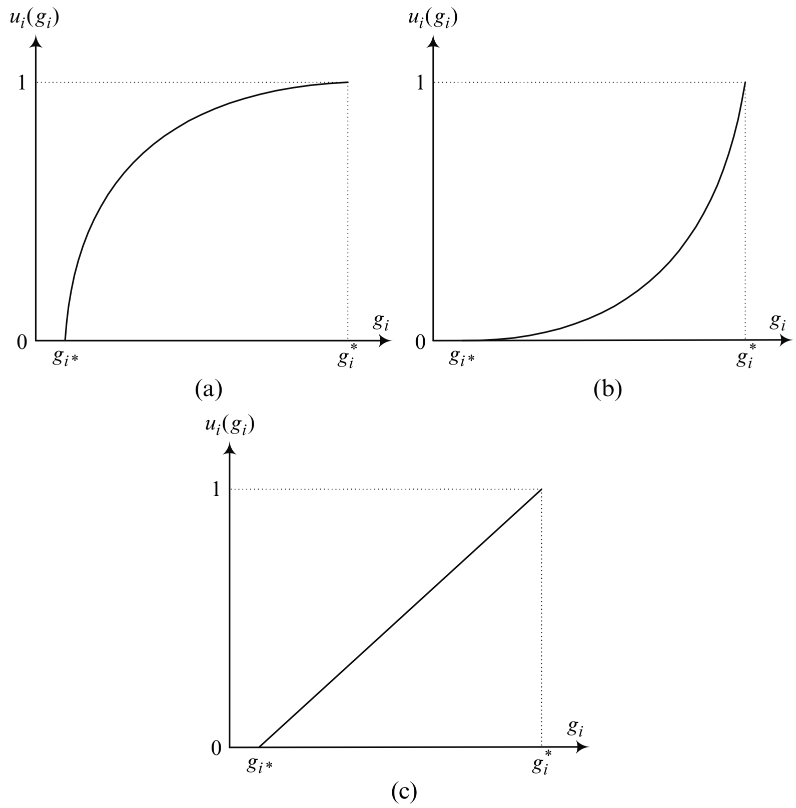

Fig. 2.1 Characteristic forms of marginal utility functions

$$g_{i*} = \max_{\forall \mathbf{x}_j \in A} \{g_{ji}\} \quad \text{and} \quad g_i^* = \min_{\forall \mathbf{x}_j \in A} \{g_{ji}\}$$

Essentially, the marginal utility functions provide a mechanism for transforming the criterion's scale into a new scale ranging in the interval $[0,1]$. This new scale represents the utility for the decision maker of each value of the criterion. The form of the marginal utility functions depends upon the decision maker's preferential system (judgment policy). Figure 2.1 presents three characteristic cases. The concave form of the utility function presented in Figure 2.1(a) indicates that the decision maker considers as quite significant small deviations from the worst performance g_i*. This corresponds with a risk-averse attitude. On the contrary, the case presented in Figure 2.1(b) corresponds with a risk-prone decision maker who is mainly interested in alternatives of top performance. Finally, the linear marginal utility function of Figure 2.1(c) indicates a risk-neutral behavior.

Transforming the criteria's scale into utility terms through the use of marginal utility functions has two major advantages:

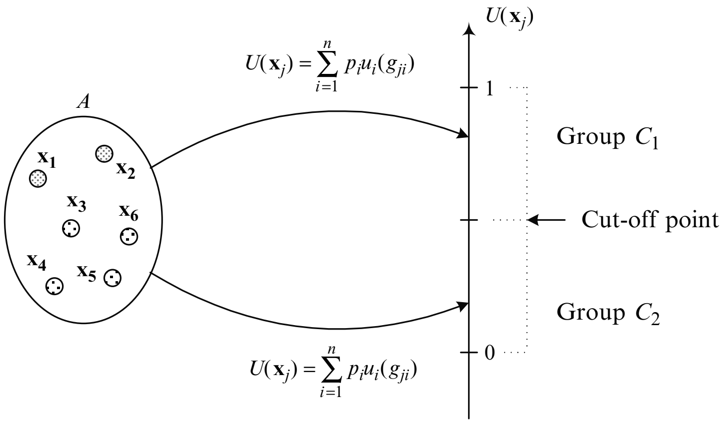

Fig. 2.2 Classification of the alternatives on the basis of their global utilities

1. It enables the modeling and representation of the nonlinear behavior of the decision maker when evaluating the performance of the alternatives.
2. It enables the consideration of qualitative criteria in a flexible way.

Given the above discussion on the concept of marginal utilities, the global utility of an alternative $\mathbf{x}_j$ specified through eq. (2.1) represents a measure of the overall performance of the alternative considering its performance on all criteria. The global utilities range in the interval [0, 1] and they constitute the criterion used to decide upon the classification of the alternatives. Figure 2.2 illustrates how the global utilities are used for classification purposes in the simple two group case. The classification is performed by comparing the global utility of each alternative with a cutoff point defined on the utility scale between 0 and 1. Alternatives with global utilities higher than the utility cutoff point are assigned into group C_1, whereas alternatives with global utilities lower than the cutoff point are assigned into group C_2.

In the general case where q groups are considered, the classification of the alternatives is performed through the following classification rules:

$$\left.\begin{array}{r} U(\mathbf{x}_j) \geq u_1 \Rightarrow \mathbf{x}_j \in C_1 \\ u_2 \leq U(\mathbf{x}_j) < u_1 \Rightarrow \mathbf{x}_j \in C_2 \\ \cdots\cdots\cdots\cdots\cdots \quad \cdots\cdots\cdots \\ U(\mathbf{x}_j) < u_{q-1} \Rightarrow \mathbf{x}_j \in C_q \end{array}\right\} \tag{2.2}$$

where $u_1, u_2, \ldots, u_{q-1}$ denote the utility cutoff points separating the group. Henceforth, these cutoff points will be referred to as utility thresholds. Essentially, each utility threshold u_k separates two consecutive groups C_k and C_{k+1}.

2.1.2 Model Development Process

2.1.2.1 General Framework

The main structural parameters of the classification model developed through the UTADIS method include the criteria weights, the marginal utility functions, and the utility thresholds. These parameters are specified through the regression-based philosophy of preference disaggregation analysis.

A general outline of the model development procedure in the UTADIS method is presented in Figure 2.3.

Initially, a reference set A' consisting of m alternatives described along n criteria is used as the training sample (henceforth the training sample will be referred to as the *reference set* in order to comply with the terminology used in MCDA). The alternatives of the reference set are classified a priori into q groups. The reference set should be constructed in such a way so that it includes an adequate number of representative examples (alternatives) from each group. Henceforth, the number of alternatives of the reference set belonging to group C_k will be denoted by m_k.

Given the classification C of the alternatives in the reference set, the objective of the UTADIS method is to develop a criteria aggregation model and a set of utility thresholds that minimize the classification error rate. The error rate refers to the differences between the estimated classification $\widehat{C}$ defined through the developed model and the prespecified classification C for the alternatives of the reference set. Such differences can be represented by introducing a binary variable E representing the classification status of each alternative:

$$E_j = \begin{cases} 0, & \text{if } \mathbf{x}_j \text{ is correctly classified} \\ 1, & \text{if } \mathbf{x}_j \text{ is misclassified} \end{cases}$$

On the basis of this binary variable, the classification error rate γ is defined as the ratio of the number of misclassified alternatives to the total number of alternatives in the reference set:

$$\gamma = \frac{\sum_{j=1}^{m} E_j}{m} \in [0, 100\%] \tag{2.3}$$

This classification error rate measure is adequate for cases where the number of alternatives of each group in the reference set is similar along all groups (i.e., $m_1 \approx m_2 \approx \cdots \approx m_q$). In the case, however, where there are significant differences, then the use of the classification error rate defined in (2.3) may lead to misleading results. For instance, consider a reference set consisting of 10 alternatives, 7 belonging into group C_1 and 3 belonging into group C_2 ($m_1 = 7$, $m_2 = 3$). In this case, a classification that assigns correctly all alternatives of group C_1 and incorrectly all alternatives of group C_2 has an error rate $\gamma = 30\%$. This is a misleading result. Actually, what should be the main point of interest in the expected classification error $\Pr(error)$. This is expressed in relation to the a priori probabilities π_1 and π_2 that an alternative belongs to groups C_1 and C_2, respectively, as follows:

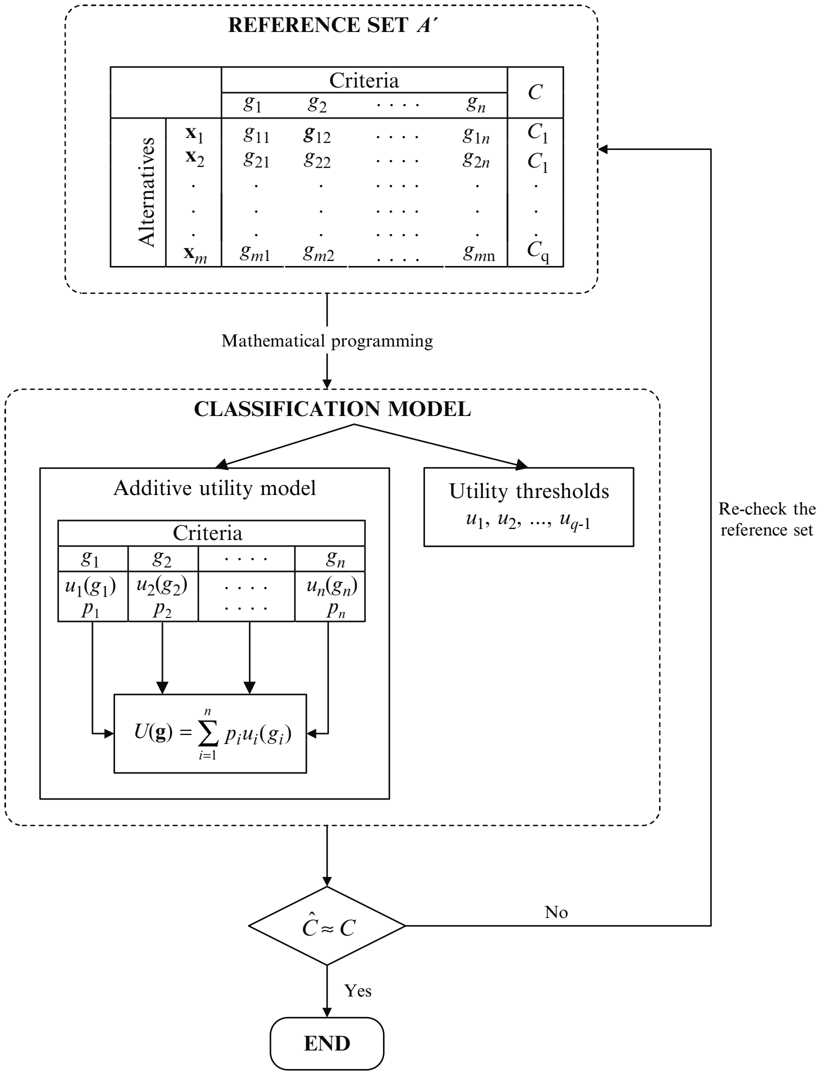

Fig. 2.3 Outline of the model development procedure in the UTADIS method

$$
\begin{aligned}
\Pr(error) &= \Pr(\text{incorrect classification of an alternative } \mathbf{x}_j) \\
&= \Pr[(\mathbf{x}_j \in C_1 \text{ and assigned in } C_2) \text{ or } (\mathbf{x}_j \in C_2 \text{ and assigned in } C_1)] \\
&= \Pr[(\mathbf{x}_j \in C_1) \wedge (\mathbf{x}_j \to C_2)] + \Pr[(\mathbf{x}_j \in C_2) \wedge (\mathbf{x}_j \to C_1)] \\
&= \pi_1 \Pr(\mathbf{x}_j \to C_2) + \pi_2 \Pr(\mathbf{x}_j \to C_1)
\end{aligned}
$$

In the above example, the error rates for the two groups (0% for C_1 and 100% for C_2) can be considered as estimates for the probabilities $\Pr(\mathbf{x}_j \to C_1)$ and $\Pr(\mathbf{x}_j \to C_2)$, respectively. Assuming that the a priori probabilities for the two groups are equal (i.e., $\pi_1 = \pi_2 = 0.5$), then the expected error of the classification is 0.5. This result indicates that the obtained classification corresponds with a random classification. In a random classification, the probabilities $\Pr(\mathbf{x}_j \to C_k)$ are determined based on the proportion of each group C_k to the total number of alternatives in the reference set. In this respect, in the above example a naïve approach would be to assign 7 out of the 10 alternatives into group C_1, i.e., $\Pr(\mathbf{x}_j \to C_1) = 0.7$, and 3 out of the 10 alternatives into group C_2, i.e., $\Pr(\mathbf{x}_j \to C_2) = 0.3$. The expected error of such a naïve approach (random classification) is 0.5.

To overcome this problem, a more appropriate measure of the expected classification error rate is expressed as follows:

$$\gamma = \sum_{k=1}^{q} \left(\pi_k \frac{\sum\limits_{\forall \mathbf{x}_j \in C_k} E_j}{m_k} \right) \in [0, 100\%] \tag{2.4}$$

Even though this measure takes into consideration the a priori probabilities of each group, it assumes that all classification errors are of equal cost to the decision maker. This is not always the case. For instance, the classification error regarding the assignment of a bankrupt firm to the group of healthy firms is much more costly than an error involving the assignment of a healthy firm to the bankrupt group. The former leads to capital cost (loss of the amount of credit granted to a firm), whereas the latter leads to opportunity cost (loss of profit that would result from granting a credit to a healthy firm). Therefore, it would be appropriate to extend the expected classification error rate (2.4) so that the costs of each individual error are also considered. The resulting measure represents the expected misclassification cost (EMC), rather than the expected classification error rate:

$$\mathrm{EMC} = \sum_{k=1}^{q} \left[\pi_k \left(\sum_{\substack{l=1 \\ l \neq k}}^{q} K_{kl} \sum_{\forall \mathbf{x}_j \in C_k} \frac{E_{klj}}{m_k} \right) \right] \in [0, 1] \tag{2.5}$$

where:

- K_{kl} is the misclassification cost involving the classification of an alternative of group C_k into group C_l ($l \neq k$).
- E_{klj} is a binary $0-1$ variable defined such that $E_{klj} = 1$ if an alternative $\mathbf{x}_j \in C_k$ is classified into group C_l ($l \neq k$) and $E_{klj} = 0$ if $\mathbf{x}_j$ is not classified into group C_l.

Comparing expressions (2.4) and (2.5), it becomes apparent that the expected classification error rate in (2.4) is a special case of the expected misclassification cost, when all costs K_{kl} are considered equal for every $k, l = 1, 2, \ldots, q$. The main difficulty related to the use of the expected misclassification cost as the appropriate measure of the quality of the obtained classification is that it is often quite difficult to have reliable estimates for the cost of each type of classification error.

If the expected classification error rate, regarding the classification of the alternatives that belong into the reference set, is considered satisfactory, then this is an indication that the developed classification model might be useful in providing reliable recommendations for the classification of other alternatives. On the other hand, if the obtained expected classification error rate indicates that the classification of the alternatives in the reference set is close to a random classification (i.e., $\gamma \approx 1/q$ or $\gamma > 1/q$), then the decision maker must check the reference set regarding its completeness and adequacy for providing representative information on the problem under consideration. Alternatively, it is also possible that the criteria aggregation model (additive utility function) is not able to provide an adequate representation of the decision maker's preferential system. In such a case, an alternative criteria aggregation model must be considered.

However, it should be pointed out that a low expected classification error rate does not necessarily ensure the practical usefulness of the developed classification model; it simply provides an indication supporting the possible usefulness of the model. On the contrary, a high expected classification error rate leads with certainty to the conclusion that the developed classification model is inadequate.

2.1.2.2 Mathematical Formulation

Pursuing the objective of the model development process in the UTADIS method, i.e., the maximization of the consistency between the estimated classification $\widehat{C}$ and the predefined one C, is performed through mathematical programming techniques.

In particular, the minimization of the expected classification error rate (2.4) requires the formulation and solution of a mixed-integer programming (MIP) problem. The solution, however, of MIP formulations is a computationally intensive procedure. Despite the significant research that has been made on the development of computationally efficient techniques for solving MIP problems within the context of classification model development, the computational effort still remains quite significant. This problem is most significant in cases where the reference set includes a large number of alternatives.

To overcome this problem, an approximation of the error rate (2.4) is used as follows:

$$\gamma' = \frac{1}{q}\sum_{k=1}^{q}\left(\frac{\sum\limits_{\forall \mathbf{x}_j \in C_k} \sigma_j}{m_k}\right) \tag{2.6}$$

where σ_j is a positive real variable, defined such that:

$$\sigma_j = \begin{cases} > 0, & \text{if } \mathbf{x}_j \text{ is misclassified} \\ 0, & \text{if } \mathbf{x}_j \text{ is classified correctly} \end{cases} \tag{2.7}$$

Essentially, σ_j represents the magnitude of the classification error for alternative $\mathbf{x}_j$. On the basis of the classification rule (2.2), the classification error for an alternative of group C_1 involves the violation of the utility threshold u_1 that defines the lower bound of group C_1. For the alternatives of the last (least preferred) group C_q, the classification error involves the violation of the utility threshold u_{q-1} that defines the upper bound of group C_q. For any other intermediate group C_k ($1 < k < q$), the

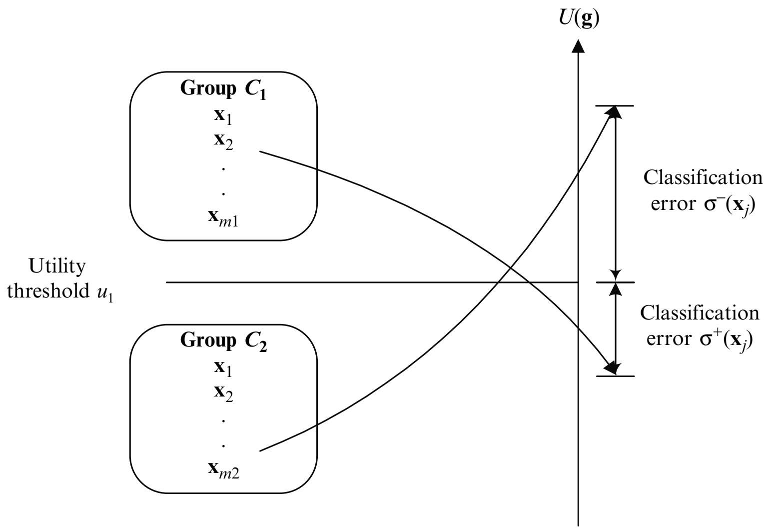

Fig. 2.4 The classification errors (two-group case)

classification error may involve either the violation of the upper bound of the group (utility threshold u_{k-1}) or the violation of the lower bound u_k.

Henceforth, the violation of the lower bound of a group will be denoted by σ^+, whereas σ^- will be used to denote the violation of the upper bound of a group. Figure 2.4 provides a graphical representation of these two errors in the simple two-group case. By definition, it is not possible that the two errors occur simultaneously (i.e., $\sigma^+\sigma^- = 0$). Therefore, the total error σ_j for an alternative $\mathbf{x}_j$ is defined as $\sigma_j = \sigma_j^+ + \sigma_j^-$.

At this point, it should be emphasized that the error functions (2.4) and (2.6) are not fully equivalent. For instance, consider a reference set consisting of four alternatives classified into two groups: $\{\mathbf{x}_1, \mathbf{x}_2\} \in C_1$, $\{\mathbf{x}_3, \mathbf{x}_4\} \in C_2$. Assume that for this reference set an additive utility classification model (CM1) is developed that misclassifies alternatives $\mathbf{x}_2$ and $\mathbf{x}_4$, such that $\sigma_2^+ = 0.2$ and $\sigma_4^- = 0.1$. Then according to (2.6) the total classification error is $\gamma' = 0.075$, whereas considering (2.4) the expected classification error rate is $\gamma = 50\%$. An alternative classification model (CM2) that classifies correctly $\mathbf{x}_2$ but retains the misclassification of $\mathbf{x}_4$ such that $\sigma_4^- = 0.5$ has $\gamma' = 0.125$ and $\gamma = 25\%$. Obviously, the model CM1 outperforms CM2 when the definition (2.6) is considered, but according to the expected classification error rate (2.4) CM2 performs better.

Despite this limitation, the definition (2.6) provides a good approximation of the expected classification error rate (2.4) while reducing the computational effort required to obtain an optimal solution.

The two forms of the classification errors can be formally expressed on the basis of the classification rule (2.2) as follows:

$$\sigma_j^+ = \max\{0, u_k - U(\mathbf{g}_j)\}, \qquad \forall \mathbf{x}_j \in C_k, k = 1, 2, \ldots, q-1$$
$$\sigma_j^- = \max\{0, U(\mathbf{g}_j) - u_{k-1}\}, \qquad \forall \mathbf{x}_j \in C_k, k = 2, 3, \ldots, q$$

These expressions illustrate better the notion of the errors. The error σ_j^+ indicates that to classify correctly a misclassified alternative $\mathbf{x}_j$ that actually belongs in group C_k, its global utility $U(\mathbf{x}_j)$ should be increased by $u_k - U(\mathbf{x}_j)$. Similarly, the σ_j^- indicates that to classify correctly a misclassified alternative $\mathbf{x}_j$ that actually belongs in C_k, its global utility $U(\mathbf{x}_j)$ should be decreased by $U(\mathbf{x}_j) - u_{k-1}$.

Introducing the error terms in the additive utility model, it is possible to rewrite the classification rule (2.2) in the form of the following constraints:

$$U(\mathbf{g}_j) + \sigma_j^+ \geq u_1, \qquad \forall \mathbf{x}_j \in C_1 \tag{2.8}$$
$$U(\mathbf{g}_j) + \sigma_j^+ \geq u_k, \qquad \forall \mathbf{x}_j \in C_k \ (k = 2, \ldots, q-1) \tag{2.9}$$
$$U(\mathbf{g}_j) - \sigma_j^- < u_{k-1}, \qquad \forall \mathbf{x}_j \in C_k \ (k = 2, \ldots, q-1) \tag{2.10}$$
$$U(\mathbf{g}_j) - \sigma_j^- < u_{q-1}, \qquad \forall \mathbf{x}_j \in C_{q-1} \tag{2.11}$$

These constraints constitute the basis for the formulation of a mathematical programming problem used to estimate the parameters of the additive utility classification model (utility thresholds, marginal utilities, criteria weights). The general form of this mathematical programming model is the following (MP):

$$\min \quad \sum_{k=1}^{q} \left[\frac{\sum_{\forall \mathbf{x}_j \in C_k} \left(\sigma_j^+ + \sigma_j^- \right)}{m_k} \right] \tag{2.12}$$

$$\text{s.t.} \quad U(\mathbf{g}_j) - u_1 + \sigma_j^+ \geq \delta_1, \qquad \forall \mathbf{x}_j \in C_1 \tag{2.13}$$
$$U(\mathbf{g}_j) - u_k + \sigma_j^+ \geq \delta_1, \qquad \forall \mathbf{x}_j \in C_k \ (k = 2, 3, \ldots, q-1) \tag{2.14}$$
$$U(\mathbf{g}_j) - u_{k-1} - \sigma_j^- \leq -\delta_2, \qquad \forall \mathbf{x}_j \in C_k \ (k = 2, 3, \ldots, q-1) \tag{2.15}$$
$$U(\mathbf{g}_j) - u_{q-1} - \sigma_j^- \leq -\delta_2, \qquad \forall \mathbf{x}_j \in C_q \tag{2.16}$$
$$U(\mathbf{g}^*) = 1 \tag{2.17}$$
$$U(\mathbf{g}_*) = 0 \tag{2.18}$$
$$u_k - u_{k+1} \geq s, \qquad k = 1, 2, \ldots, q-1 \tag{2.19}$$
$$u_i(g_i) \text{ increasing functions} \tag{2.20}$$
$$\sigma_j^+, \sigma_j^- \geq 0, \qquad j = 1, 2, \ldots, m \tag{2.21}$$

In constraints (2.13)–(2.14), δ_1 is a positive constant used to avoid cases where $U(\mathbf{g}_j) = u_k$ when $\mathbf{x}_j \in C_k$. Of course, u_k is considered as the lower bound of group C_k. In this regard, the case $\delta_1 = 0$, typically, does not pose any problem during model development and implementation. However, assuming the simple two-group case, the specification $\delta_1 = 0$ may lead to the development of a classification model for which $U(\mathbf{g}_j) = u_1 = 1$, for all $\mathbf{x}_j \in C_1$, and $U(\mathbf{g}_j) < u_1 = 1$ for all $\mathbf{x}_j \in C_2$. Because the utility threshold u_1 is defined as the lower bound of group C_1, it is obvious that such a model performs an accurate classification of the alternatives. Practically, however, because all alternatives of group C_1 are placed on the utility threshold, the generalizing ability of such a model is expected to be limited. Therefore, to avoid such situations, a small positive (non-zero) value for the constant δ_1 should be chosen. The constant δ_2 in (2.15)–(2.16) is used in a similar way.

Constraints (2.17) and (2.18) are used to normalize the global utilities in the interval $[0,1]$. In these constraints, $\mathbf{g}_*$ and $\mathbf{g}^*$ denote the vectors consisting of the least and the most preferred levels of the evaluation criteria. Finally, constraint (2.19) is used to ensure that the utility threshold u_k is higher than the utility threshold u_{k+1}, thus ensuring the ordering of the groups from the most preferred (C_1) to the least preferred ones (C_q). In this ordering of the groups, higher utilities are assigned to the most preferred groups. In constraint (2.19), s is a constant defined such that $s > \delta_1, \delta_2$.

Introducing the additive utility function (2.1) in MP leads to the formulation of a nonlinear programming problem. This is because the additive utility function (2.1) has two unknown parameters: (a) the criteria weights and (b) the marginal utility functions. Therefore, constraints (2.13)–(2.18) take a nonlinear form, and the solution of the resulting nonlinear programming problem can be cumbersome. To overcome this problem, the additive utility function (2.1) is rewritten in a simplified form as follows:

$$U(\mathbf{g}) = \sum_{i=1}^{n} u'_i(g_i) \tag{2.22}$$

where:

$$\left.\begin{array}{l} u'_i(g_i) = p_i u_i(g_i) \\ u'_i(g_{i*}) = 0 \\ u'_i(g_i^*) = p_i \end{array}\right\} \tag{2.23}$$

Both (2.1) and (2.22) are equivalent expressions for the additive utility function. Nevertheless, the latter requires only the specification of the marginal utility functions $u'_i(g_i) \in [0, p_i]$. As illustrated in Figure 2.1, these functions can be of any form. The UTADIS method does not prespecify a functional form for these functions. Therefore, it is necessary to express the marginal utility functions in terms of specific decision variables to be estimated through the solution of MP. This is achieved through the modeling of the marginal utilities as piece-wise linear functions through a process that is graphically illustrated in Figure 2.5.

The range $[g_{i*}, g_i^*]$ of each criterion is divided into $a_i - 1$ subintervals $[g_i^h, g_i^{h+1}]$, $h = 1, 2, \ldots, a_i - 1$. The estimation of the unknown marginal utility functions can be performed by estimating the marginal utilities at the break-points $g_i^2, \ldots, g_i^{a_i}$. As illustrated in Figure 2.5, this estimation provides an approximation of the true marginal utility functions. On the basis of this approach, it would be reasonable to assume that the larger the number of subintervals that are specified, the better is the approximation of the marginal utility functions. The definition of a large number of subintervals, however, provides increased degrees of freedom to the additive utility model. This increases the fitting ability of the developed model to the data of

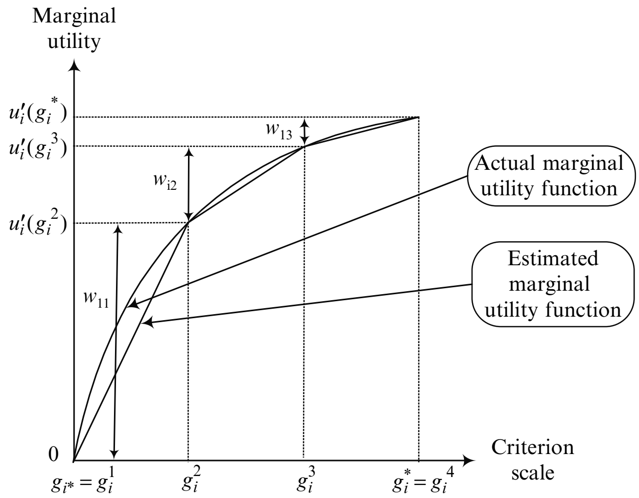

Fig. 2.5 Piece-wise linear form of marginal utility functions

the reference set; the instability, however, of the model is also increased (the model becomes sample-based).

The marginal utility at the break-point g_i^h is written as follows:

$$u_i'(g_i^h) = \sum_{t=1}^{h-1} w_{it}$$

where $w_{it} = u'(g_i^t) - u'(g_i^{t-1}) \geq 0$ are the parameters that must be estimated in order to specify the marginal value function. With this modeling, the marginal value function of any alternative $\mathbf{x}_j$ on the criterion g_i is expressed as follows:

$$u'(g_{ji}) = \sum_{t=1}^{r_{ji}-1} w_{it} + \frac{g_{ji} - g_i^{r_{ji}}}{g_i^{r_{ji}+1} - g_i^{r_{ji}}} w_{i,r_{ji}}$$

where r_{ji} $(1 \leq r_{ji} \leq a_i - 1)$ denotes the subinterval $[g_i^{r_{ji}},\ g_i^{r_{ji}+1}]$ into which the performance g_{ji} of alternative $\mathbf{x}_j$ on criterion g_i belongs to. The global utility of the alternative $\mathbf{x}_j$ is also expressed in terms of the unknown parameters w:

$$U(\mathbf{g}_j) = \sum_{i=1}^{n} \left(\sum_{t=1}^{r_{ji}-1} w_{it} + \frac{g_{ji} - g_i^{r_{ji}}}{g_i^{r_{ji}+1} - g_i^{r_{ji}}} w_{i,r_{ji}} \right)$$

Therefore, the problem MP is explicitly written as the following linear programming problem (LP):

$$\min \sum_{k=1}^{q} \left[\frac{\sum\limits_{\forall \mathbf{x}_j \in C_k} \left(\sigma_j^+ + \sigma_j^- \right)}{m_k} \right] \tag{2.24}$$

$$\text{s.t.} \sum_{i=1}^{n} \left(\sum_{t=1}^{r_{ji}-1} w_{it} + \frac{g_{ji} - g_i^{r_{ji}}}{g_i^{r_{ji}+1} - g_i^{r_{ji}}} w_{i,r_{ji}} \right) - u_1 + \sigma_j^+ \geq \delta_1, \qquad \forall \mathbf{x}_j \in C_1 \tag{2.25}$$

$$\sum_{i=1}^{n} \left(\sum_{t=1}^{r_{ji}-1} w_{it} + \frac{g_{ji} - g_i^{r_{ji}}}{g_i^{r_{ji}+1} - g_i^{r_{ji}}} w_{i,r_{ji}} \right) - u_k + \sigma_j^+ \geq \delta_1, \qquad \forall \mathbf{x}_j \in \{C_2, \ldots, C_{q-1}\} \tag{2.26}$$

$$\sum_{i=1}^{n} \left(\sum_{t=1}^{r_{ji}-1} w_{it} + \frac{g_{ji} - g_i^{r_{ji}}}{g_i^{r_{ji}+1} - g_i^{r_{ji}}} w_{i,r_{ji}} \right) - u_{k-1} - \sigma_j^- \leq -\delta_2, \quad \forall \mathbf{x}_j \in \{C_2, \ldots, C_{q-1}\} \tag{2.27}$$

$$\sum_{i=1}^{n} \left(\sum_{t=1}^{r_{ji}-1} w_{it} + \frac{g_{ji} - g_i^{r_{ji}}}{g_i^{r_{ji}+1} - g_i^{r_{ji}}} w_{i,r_{ji}} \right) - u_{q-1} - \sigma_j^- \leq -\delta_2, \quad \forall \mathbf{x}_j \in C_q \tag{2.28}$$

$$\sum_{i=1}^{n} \sum_{t=1}^{a_i-1} w_{it} = 1 \tag{2.29}$$

$$u_k - u_{k-1} \geq s, \qquad 1 \leq k \leq q-1 \tag{2.30}$$

$$w_{it}, \sigma_j^+, \sigma_j^- \geq 0, \qquad \forall j, i, t \tag{2.31}$$

Constraints (2.25)–(2.28), and (2.29)–(2.30) correspond with the constraints (2.13)–(2.16), (2.17), and (2.19) of MP. The non-negativity constraint on the variables w ensures that the marginal value functions are increasing (constraint (2.20) in MP).

2.1.3 Model Development Issues

The simple linear form of LP ensures the existence of a global optimum solution. However, often there are multiple optimal solutions. The existence of multiple optimal solutions is most often when the groups are perfectly separable, i.e., when there is no group overlap. In such cases, all error variables σ_j^+ and σ_j^- are zero. The determination of a large number of criteria subintervals is positively related to the existence of multiple optimal solutions (as already mentioned, as the number of subintervals increases, the degrees of freedom of the developed additive utility model also increases and so does the fitting ability of the model).

In addition to the above phenomenon, it is also important to emphasize that even if a unique optimal solution does exist for LP, its stability needs to be carefully considered. A solution is considered to be stable if it is not significantly affected by small trade-offs to the objective function (i.e., if near-optimal solutions are quite similar to the optimal one). The instability of the optimal solution is actually the result of overfitting the developed additive utility model to the alternatives of the reference set. This may affect negatively the generalizing classification performance of the developed classification model. In addition to the classification performance issue, the instability of the additive utility model also raises interpretation problems. If the developed model is unstable, then it is clearly very difficult to derive secure conclusions on the contribution of the criteria in the classification of the alternatives (the criteria weights are unstable and therefore difficult to interpret).

The consideration of these issues in the UTADIS method is performed through a post-optimality analysis that follows the solution of LP. The objective of post-optimality analysis is to explore the existence of alternate optimal solutions and near-optimal solutions. There are many different ways that can be used to perform the post-optimality stage considering the parameters that are involved in the model development process. These parameters include the constants δ_1, δ_2, and s, as well as the number of criteria subintervals. The use of mathematical programming techniques provides increased flexibility in considering a variety of different forms for the post-optimality analysis. Some issues that are worth the consideration in the post-optimality stage include:

1. The maximization of the constants δ_1 and δ_2. This implies a maximization of the minimum distance between the correctly classified alternatives and the utility thresholds, thus resulting in a more clear separation of the groups.
2. Maximization of the sum of the differences between the global utilities of the correctly classified alternatives from the utility thresholds. This approach extends the previous point considering all differences instead of the minimum ones.
3. Minimization of the total number of misclassified alternatives using the error function (2.4).
4. Determination of the minimum number of criteria subintervals.

Considering, however, the issues regarding the stability of the developed model and its interpretation, none of these approaches ensures the existence of a unique and stable solution. Consequently, the uncertainty on the interpretation of the model is still an issue to be considered.

To overcome this problem, the post-optimality stage performed in the UTADIS method focuses on the investigation of the stability of the criteria weights rather than on the consideration of the technical parameters of the model development process. In particular, during the post-optimality stage $n+q-1$ new linear programs are solved, each having the same form with LP. The solution of LP2 is used as input to each of these new linear programs to explore the existence of other optimal or near-optimal solutions. The objective function of each problem s involves the maximization of each criterion weight (for $s = 1, 2, \ldots, n$) and the value of the utility thresholds (for $s > n$). All new solutions found during the post-optimality stage are optimal or near optimal for LP. This is ensured by imposing the following constraint:

$$f' \leq (1+z) f^*$$

where:

- f^* is the optimal value for the objective function of LP,
- f' is the value of the objective function of LP evaluated for any new solution obtained during the post-optimality stage.
- z is a small portion of f^* (a trade-off made to the optimal value of the objective function in order to investigate the existence of near-optimal solutions).

This constraint in added to the formulation of LP, and the new linear program that is formed is solved to maximize either the criteria weights or the utility thresholds as noted above. Finally, the additive utility model used to perform the classification of the alternatives is formed from the average of all solutions obtained during the post-optimality stage.

Overall, despite the problems raised by the existence of multiple optimal solutions, it should be noted that LP provides consistent estimates for the parameters of the additive utility classification model. The consistency property for mathematical programming formulations used to estimate the parameters of a decision-making model was first introduced by Charnes et al. (1955). The authors consider a mathematical programming formulation to satisfy the consistency property if it provides estimates of the model's parameters that approximate (asymptotically) the true values of the parameters as the number of observations (alternatives) used for model developed increases. According to the authors, this is the most significant property that a mathematical programming formulation used for model development should have, as it ensures that the formulation is able to identify the true values of the parameters under consideration, given that enough information is available.

LP has the consistency property. Indeed, as new alternatives are added in an existing reference set and given that these alternatives add new information (i.e., they are not dominated by alternatives already belonging in the reference set), then the new alternatives will add new non-redundant constraints in LP. These constraints reduce the size of the feasible set. Asymptotically, for large reference sets, this will lead to the identification of a unique optimal solution that represents the decision-maker's judgment policy and preferential system.

2.2 The Multigroup Hierarchical Discrimination Method (MHDIS)

2.2.1 Outline and Main Characteristics

People often employ, sometimes intuitively, a sequential/hierarchical process to classify alternatives to groups using available information and holistic judgments. For example, examine if an alternative can be assigned to the best group C_1, if not then try the second-best group C_2, etc. This is the logic of the MHDIS method and (Zopounidis and Doumpos, 2000b) its main distinctive feature compared with the UTADIS method. A second major difference between the two methods involves the mathematical programming framework used to develop the classification models.

Model development in UTADIS is based on a linear programming formulation followed by a post-optimality stage. In MHDIS, the model development process is performed using two linear programs and a mixed integer one that gradually calibrate the developed model so that it accommodates two objectives: (1) the minimization of the total number of misclassifications, and (2) the maximization of the clarity of the classification. These two objectives are pursued through a lexicographic approach, i.e., initially the minimization of the total number of misclassifications is sought and then the maximization of the clarity of the classification is performed. The common feature shared by both MHDIS and UTADIS involves the form of the criteria aggregation model that is used to model the decision-maker's preferences in classification problems, i.e., both methods employ a utility-based framework.

2.2.2 The Hierarchical Discrimination Process

The MHDIS method proceeds progressively in the classification of the alternatives into the predefined groups. The hierarchical discrimination process used in MHDIS consists of $q-1$ stages (Figure 2.6). Each stage k is considered as a two-group classification problem, where the objective is to discriminate the alternatives of group C_k from the alternatives of the other groups. Because the groups are defined in an ordinal way, this is translated to the discrimination of group C_k from the set of groups $\{C_{k+1}, C_{k+2}, \ldots, C_q\}$. Therefore at each stage of the hierarchical discrimination process, two choices are available for the classification of an alternative:

1. To decide that the alternative belongs in group C_k, or
2. To decide that the alternative belongs at most in the group C_{k+1} (i.e., it belongs in one of the groups C_{k+1} to C_q).

Within this framework, the procedure starts from group C_1 (most preferred alternatives). The alternatives found to belong in group C_1 (correctly or incorrectly) are excluded from further consideration. In a second stage, the objective is to identify the alternatives belonging in group C_2. Once again, all the alternatives found to belong in this group (correctly or incorrectly) are excluded from further consideration, and the same procedure continues until all alternatives are classified into the predefined groups.

The criteria aggregation model used to decide upon the classification of the alternatives at each stage k of the hierarchical discrimination process has the form of an additive utility function, similar to the one used in UTADIS.

$$U_k(\mathbf{g}) = \sum_{i=1}^{n} u_{ki}(g_i) \in [0,1] \tag{2.32}$$

$U(\mathbf{g})$ denotes the utility of classifying any alternative into group C_k on the basis of the alternative's performance on the set of criteria $\mathbf{g}$, and $u_{ki}(g_i)$ denotes the corresponding marginal utility function regarding the classification of any alternative

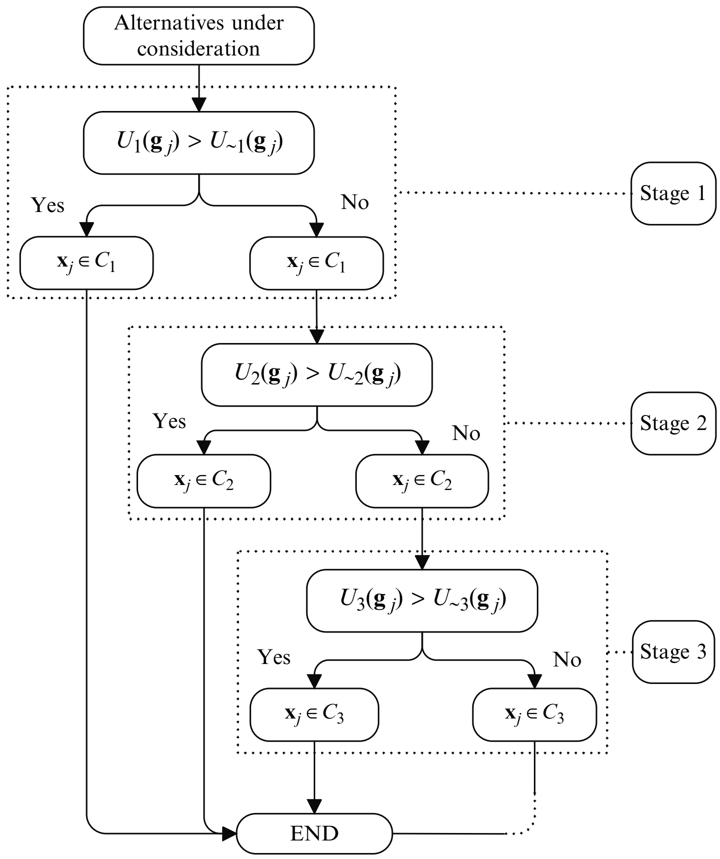

Fig. 2.6 The hierarchical discrimination process in MHDIS

into group C_k according to a specific criterion g_i. Conceptually, the utility function $U_k(\mathbf{g})$ provides a measure of the similarity of the alternatives to the characteristics of group C_k.

Nevertheless, as noted above, at each stage k of the hierarchical discrimination process there are two choices available for the classification of an alternative, the classification into group C_k and the classification at most into group C_{k+1}. The utility function $U_k(\mathbf{g})$ measures the utility (value) of the first choice. To make a classification decision, the utility of the second choice (i.e., classification at most into group C_{k+1}) needs also to be considered. This is measured by a second utility function denoted by $U_{\sim k}(\mathbf{g})$ that has the same form (2.32).

Based on these two utility functions, the classification of an alternative $\mathbf{x}_j$ is performed using the following rules:

$$\left.\begin{array}{l}\text{if } U_k(\mathbf{g}_j) > U_{\sim k}(\mathbf{g}_j) \text{ then } \mathbf{x}_j \in C_k \\ \text{if } U_k(\mathbf{g}_j) < U_{\sim k}(\mathbf{g}_j) \text{ then } \mathbf{x}_j \in C_k^{>}\end{array}\right\} \tag{2.33}$$

where $C_k^{>}$ denotes the set of groups $\{C_{k+1}, C_{k+2}, \ldots, C_q\}$. During model development, the case $U_k(\mathbf{g}_j) = U_{\sim k}(\mathbf{g}_j)$ is considered to be a misclassification. When the developed additive utility functions are used for extrapolating purposes, such a case indicates that the classification of the alternatives is not clear and additional analysis is required. This analysis can be based on the examination of the marginal utilities $u_{ki}(g_{ji})$ and $u_{\sim ki}(g_{ji})$ to determine how the performance of the alternatives on each of the evaluation criterion affects their classification.

In both utility functions $U_k(\mathbf{g})$ and $U_{\sim k}(\mathbf{g})$, the corresponding marginal utilities $u_{ki}(g_{ji})$ and $u_{\sim ki}(g_{ji})$ are monotone functions on the criteria scale. The marginal utility functions $u_{ki}(g_{ji})$ are increasing, whereas $u_{\sim ki}(g_{ji})$ are decreasing functions. This specification is based on the ordinal definition of the groups. In particular, because the alternatives of group C_k are considered to be preferred to the alternatives of the groups C_{k+1} to C_q, it is expected that the higher the performance of an alternative on criterion g_i, the more similar the alternative is to the characteristics of group C_k (increasing form of the marginal utility function $u_{ki}(g_{ji})$) and the less similar it is to the characteristics of the groups C_{k+1} to C_q (decreasing form of the marginal utility function $u_{\sim ki}(g_{ji})$).

The marginal utility functions are modeled in a piece-wise linear form, similar to the case of the UTADIS method. The piece-wise linear modeling of the marginal utility functions in the MHDIS method is illustrated in Figure 2.7. In contrast with the UTADIS method, the criteria's scale is not divided into subintervals. Instead, the performance of each reference alternative is considered as a distinct criterion level. For instance, assuming that the reference set includes m alternatives each having a different performance on criterion g_i, then m criterion levels are considered, ordered from the least preferred one $g_{i*} = \min\{g_{ji}\}$, $\forall \mathbf{x}_j \in A$ to the most preferred one $g_i^* = \max\{g_{ji}\}$, $\forall \mathbf{x}_j \in A$, where a_i is the number of unique values for criterion g_i (e.g., in this example $a_i = m$). Denoting as g_i^h and g_i^{h+1} two consecutive levels of criterion g_i ($g_i^{h+1} > g_i^h$), the monotonicity of the marginal utilities is imposed through the following constraints (z is a small positive constant):

$$w_{kih} \geq z \quad \text{and} \quad w_{\sim kih} \geq z$$

where,

$$\begin{aligned} w_{kih} &= u_{ki}(g_i^{h+1}) - u_{ki}(g_i^h) \\ w_{\sim kih} &= u_{\sim ki}(g_i^h) - u_{\sim ki}(g_i^{h+1}) \end{aligned}$$

Thus, it is possible to express the global utility of an alternative $\mathbf{x}_j$ in terms of the incremental variables w as follows:

$$U_k(\mathbf{g}_j) = \sum_{i=1}^{n} \sum_{h=1}^{r_{ji}-1} w_{kih} \quad \text{and} \quad U_{\sim k}(\mathbf{g}_j) = \sum_{i=1}^{n} \sum_{h=r_{ji}}^{a_i-1} w_{\sim kij} \tag{2.34}$$

Although both UTADIS and MHDIS employ a utility-based modeling framework, it should be emphasized that the marginal utility functions in MHDIS do not

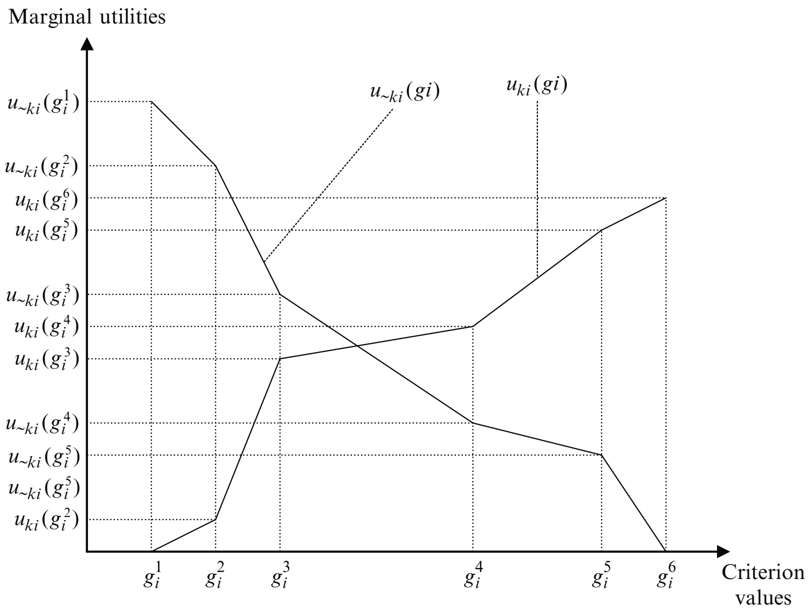

Fig. 2.7 Piece-wise linear form of the marginal utility functions in MHDIS

indicate the performance of an alternative with regard to an evaluation criterion; they rather serve as a measure of the conditional similarity of an alternative $\mathbf{x}_j$ to the characteristics of group C_k (on the basis of a specific criterion) when the choice among C_k and all the lower (worse) groups $C_{k+1},\ldots,C_q$ is considered. In this regard, a high marginal utility $u_{ki}(g_{ji})$ would indicate that when considering the performance of alternative $\mathbf{x}_j$ on criterion g_i, the most appropriate decision would be to assign the alternative into group C_k instead of the set of groups $\{C_{k+1},\ldots,C_q\}$ (the overall classification decision depends upon the examination of all criteria). This simple example indicates that the use of utilities in MHDIS does not correspond to the alternatives themselves, but rather to the appropriateness of the choices (classification decisions) that the decision maker has measured on the basis of the alternatives' performances on the evaluation criteria.

2.2.3 Estimation of Utility Functions

According to the hierarchical discrimination procedure described above, the classification of the alternatives in q classes requires the development of $2(q-1)$ utility functions. The estimation of these utility functions in MHDIS is accomplished through mathematical programming techniques. In particular, at each stage of the hierarchical discrimination procedure, two linear programs and a mixed-integer one are solved to estimate "optimally" both utility functions.[1] The term "optimally" refers to the classification of the alternatives of the reference set, such that (1) the total number of misclassifications is minimized and (2) the clarity of the classification is maximal.

These two objectives are addressed lexicographically through the sequential solution of two linear programming problems (LP1 and LP2) and a mixed-integer programming problem (MIP). Essentially, the rationale behind the sequential solution of these mathematical programming problems is the following. As noted in the discussion of the UTADIS method, the direct minimization of the total classification error (cf. equations (2.4) or (2.5)) is a quite complex and hard problem to face, from a computational effort point of view. To cope with this problem in UTADIS, an approximation was introduced (cf. equation (2.6)) considering the magnitude of the violations of the classification rules, rather than the number of violations, which defines the classification error rate. As noted, this approximation overcomes the problem involving the computational intensity of optimizing the classification error rate. Nevertheless, the results obtained from this new error function are not necessarily optimal when the classification error rate is considered. To address these issues, MHDIS combines the error function (2.6) with the actual classification error rate. In particular, initially an error function of the form of (2.6) is employed to identify the alternatives of the reference set that are hard to classify correctly (i.e., they are misclassified). This is performed through a linear programming formulation (LP1). Generally, the number of these alternatives is expected to be a small portion of the number of alternatives in the reference set. Then, a more direct error minimization approach is used considering only this reduced set of misclassified alternatives. This approach considers the actual classification error (2.4). The fact that the analysis at this stage focuses only a reduced part of the reference set (i.e., the misclassified alternatives) significantly reduces the computational effort required to minimize the actual classification error function (2.4). The minimization of this error function is performed through a MIP formulation. Finally, given the optimal classification model obtained through the solution of MIP, a linear programming formulation (LP2) is employed to maximize the clarity of the obtained classification without changing the groups into which the alternatives are assigned. The details of this three-step process are described below, along with the mathematical programming formulations used at each step.

LP1: Minimizing the Overall Classification Error

The initial step in the model development process is based on a linear programming formulation. In this formulation, the classification errors are considered as

[1] Henceforth, the discussion focuses on the development of a pair of utility functions at stage k of the hierarchical discrimination process. The first utility function $U_k(\mathbf{g})$ characterizes the alternatives of group C_k, whereas the second utility function $U_{\sim k}(\mathbf{g})$ characterizes the alternatives belonging in the set of groups $\{C_{k+1}, C_{k+2}, \ldots, C_q\}$. The same process applies to all stages $k = 1, 2, \ldots, q-1$ of the hierarchical discrimination process.

real-valued variables, defined similar to the error variables σ^+ and σ^- used in the UTADIS method. In the case of the MHDIS method, these error variables are defined through the classification rule (2.33):

$$\sigma_{kj}^{+} = \max\{0, U_{\sim k}(\mathbf{g}_j) - U_k(\mathbf{g}_j)\}, \qquad \forall \mathbf{x}_j \in C_k$$
$$\sigma_{kj}^{-} = \max\{0, U_k(\mathbf{g}_j) - U_{\sim k}(\mathbf{g}_j)\}, \qquad \forall \mathbf{x}_j \in C_k^{>}$$

Essentially, the error σ^+ indicates the misclassification of an alternative toward a lower (worst) group compared with the one where it actually belongs, whereas the error σ^- indicates a misclassification toward a higher (better) group. Both errors refer to a specific stage k of the hierarchical model development process.

On the basis of the above considerations, the initial linear program (LP1) to be solved is the following:

$$\min \quad \sum_{k=1}^{q} \left[\frac{\sum_{\forall \mathbf{x}_j \in C_k} \left(\sigma_{kj}^{+} + \sigma_{kj}^{-} \right)}{m_k} \right] \tag{2.35}$$

$$\text{s.t.} \quad \sum_{i=1}^{n} \sum_{h=1}^{r_{ji}-1} w_{kih} - \sum_{i=1}^{n} \sum_{h=r_{ji}}^{a_i-1} w_{\sim kih} + \sigma_{kj}^{+} \geq s, \qquad \forall \mathbf{x}_j \in C_k \tag{2.36}$$

$$\sum_{i=1}^{n} \sum_{j=r_{ji}}^{a_i-1} w_{\sim kih} - \sum_{i=1}^{m} \sum_{j=1}^{r_{ji}-1} w_{kih} + \sigma_{kj}^{-} \geq s, \qquad \forall \mathbf{x}_j \in C_k^{>} \tag{2.37}$$

$$w_{kih} \geq z, w_{\sim kih} \geq z \tag{2.38}$$

$$\sum_{i=1}^{m} \sum_{j=1}^{a_i-1} w_{kij} = 1, \sum_{i=1}^{m} \sum_{j=1}^{a_i-1} w_{\sim kij} = 1 \tag{2.39}$$

$$\sigma_{kj}^{+}, \sigma_{kj}^{-} \geq 0 \tag{2.40}$$

s, t small positive constants

Constraints (2.36)–(2.37) define the classification error variables σ_{kj}^{+} and σ_{kj}^{-}. These constraints are formulated on the basis of the classification rule (2.33) and the global utility functions (2.34). In the right-hand side of these constraints, a small positive constant s is used to impose the inequalities of the classification rule (2.33). This constant is similar to the constants δ_1 and δ_2 used in the linear programming formulation of the UTADIS method. The set of constraints defined in (2.38) is used to ensure the monotonicity of the marginal utility functions, whereas the set of constraints in (2.39) normalize the global utility to range between 0 and 1.

MIP: Minimizing the number of misclassifications

The solution of LP1 leads to the development of an initial pair of utility functions $U_k(\mathbf{g})$ and $U_{\sim k}(\mathbf{g})$ that discriminate group C_k from the groups C_{k+1} to C_q. These utility functions define a classification of the alternatives in the reference set that is optimal considering the classification error measured in terms of the real-valued variables σ_{kj}^{+} and σ_{kj}^{-}. When the classification error rate is considered, however, these utility functions may lead to suboptimal results. Nevertheless, this initial pair of utility functions enables the identification of the alternatives that can be easily classified correctly and the "hard" alternatives. The "hard" alternatives are the ones misclassified by the pair of utility functions developed through the solution of LP1. Henceforth, the set of alternatives classified correctly by LP1 will be denoted by *COR*, whereas the set of misclassified alternatives will be denoted by *MIS*.

Assuming that the set *MIS* includes at least two alternatives, it is possible to achieve a "rearrangement" of the magnitude of the classification errors σ_{kj}^{+} and σ_{kj}^{-} for the misclassified alternatives (alternatives of *MIS*) that will lead to the reduction of the number of misclassifications. However, as it has already been noted, this requires the introduction of binary 0-1 error variables to MIP model. To avoid the increased computational effort required to solve MIP problems, the MIP formulation used in MHDIS considers only the misclassifications that occur through the solution of LP1, while retaining all the correct classifications. Thus, it becomes apparent that actually, LP1 is an exploratory problem whose output is used as input information to MIP. This reduces significantly the number of binary 0-1 variables, which are associated with each misclassified alternative, thus alleviating the computational effort required to obtain a solution.

While this sequential consideration of LP1 and MIP considerably reduces the computational effort required to minimize the classification error rate, it should be emphasized that the obtained classification model may be near optimal instead of globally optimal. This is due to the fact that MIP inherits the solution of LP1. Therefore, the number of misclassifications attained after solving MIP depends on the optimal solution identified by LP1 (i.e., different optimal solutions of LP1 may lead to different number of misclassifications by MIP). Nevertheless, using LP1 as a preprocessing stage to provide an input to MIP provides an efficient mechanism (in terms of computational effort) to obtain an approximation of the globally minimum number of misclassifications. Formally, MIP is expressed as follows:

$$\min \quad \sum_{k=1}^{q} \left[\frac{\sum_{\forall \mathbf{x}_j \in C_k \cap MIS} \left(E_{kj}^{+} + E_{kj}^{-} \right)}{m'_k} \right] \tag{2.41}$$

$$\text{s.t.} \quad \sum_{i=1}^{n} \sum_{j=1}^{r_{ji}-1} w_{kih} - \sum_{i=1}^{n} \sum_{h=r_{ji}}^{a_i-1} w_{\sim kih} \geq s, \qquad \forall \mathbf{x}_j \in C_k \cap COR \tag{2.42}$$

$$\sum_{i=1}^{n}\sum_{h=r_{ji}}^{a_i-1} w_{\sim kih} - \sum_{i=1}^{n}\sum_{j=1}^{r_{ji}-1} w_{kih} \geq s, \qquad \forall \mathbf{x}_j \in C_k^{>} \cap COR \tag{2.43}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{r_{ji}-1} w_{kih} - \sum_{i=1}^{n}\sum_{h=r_{ji}}^{a_i-1} w_{\sim kih} + E_{kj}^{+} \geq s \qquad \forall \mathbf{x}_j \in C_k \cap MIS \tag{2.44}$$

$$\sum_{i=1}^{n}\sum_{h=r_{ji}}^{a_i-1} w_{\sim kih} - \sum_{i=1}^{n}\sum_{j=1}^{r_{ji}-1} w_{kih} + E_{kj}^{-} \geq s, \qquad \forall \mathbf{x}_j \in C_k^{>} \cap MIS \tag{2.45}$$

$$w_{kih} \geq z, w_{\sim kih} \geq z \tag{2.46}$$

$$\sum_{i=1}^{m}\sum_{j=1}^{a_i-1} w_{kij} = 1, \ \sum_{i=1}^{m}\sum_{j=1}^{a_i-1} w_{\sim kij} = 1 \tag{2.47}$$

$$E_{kj}^{+}, E_{kj}^{-} \in \{0,1\} \tag{2.48}$$

s, t small positive constants

Constraints (2.42) and (2.43) are used to ensure that all correct classifications achieved by solving LP1 are retained. Constraints (2.44)–(2.45) are used only for the alternatives that were misclassified by LP1 (set *MIS*). Their interpretation is similar to the constraints (2.36) and (2.37) in LP1. Their only difference is the transformation of the real-valued error variables σ^+ and σ^- of LP1 into the binary 0-1 variables E^+ and E^- that indicate the classification status of an alternative. Constraints (2.44)–(2.45) define these binary variables as follows: $E_{kj}^{+} = 1$ indicates that the alternative $\mathbf{x}_j$ of group C_k is classified by the developed model into the set of groups $C_k^{>}$, whereas $E_{kj}^{-} = 1$ indicates that the alternative $\mathbf{x}_j$ belonging in one of the groups C_{k+1} to C_q ($C_k^{>}$) is classified by the developed model into group C_k. Both cases are misclassifications. On the contrary, the cases $E_{kj}^{+} = 0$ and $E_{kj}^{-} = 0$ indicate the correct classification of the alternative $\mathbf{x}_j$. The interpretation of constraints (2.46) and (2.47) has already been discussed for the LP1 formulation. The objective of MIP involves the minimization of a weighted sum of the error variables E^+ and E^-. The weighting is performed considering the number of alternatives in the set *MIS* from each group C_k. This is denoted by m_k'.

LP2: Maximizing the minimum distance

Solving LP1 and then MIP leads to the "optimal" classification of the alternatives, where the term "optimal" refers to the minimization of the number of misclassified alternatives. However, it is possible that the correct classification of some alternatives is "marginal." This situation appears when the classification rules (2.33) are marginally satisfied, i.e., when there is only a slight difference between $U_k(\mathbf{g}_j)$ and $U_{\sim k}(\mathbf{g}_j)$. For instance, assume a pair of utility functions developed such that for an alternative $\mathbf{x}_j$ of group C_k, its global utilities are $U_k(\mathbf{g}_j) = 0.5$ and $U_{\sim k}(\mathbf{g}_j) = 0.498$. Given these utilities and considering the classification rules (2.33), it is obvious that alternative $\mathbf{x}_j$ is classified in the correct group (i.e., in group C_k). This is, however, a marginal result. Instead, another pair of utility functions for which $U_k(\mathbf{g}_j) = 0.8$ and $U_{\sim k}(\mathbf{g}_j) = 0.1$ is clearly preferred, providing a more clear conclusion.

This issue is addressed in MHDIS through a third mathematical programming formulation used on the basis of the optimal solution of MIP. At this stage the minimum difference d between the global utilities of the correctly classified alternatives identified after solving MIP is introduced:

$$d = \min\left\{ \min_{\mathbf{x}_j \in C_k \cap COR'} \left\{ U_k(\mathbf{g}_j) - U_{\sim k}(\mathbf{g}_j) \right\},\ \min_{\mathbf{x}_j \in C_k^{>} \cap COR'} \left\{ U_{\sim k}(\mathbf{g}_j) - U_k(\mathbf{g}_j) \right\} \right\}$$

where COR' denotes the set of alternatives classified correctly by the pair of utility functions developed through the solution of MIP. The objective of this third phase of the model development procedure is to maximize d. This is performed through the following linear programming formulation (LP2).

$$\min \quad d \tag{2.49}$$

$$\text{s.t.} \quad \sum_{i=1}^{n} \sum_{j=1}^{r_{ji}-1} w_{kih} - \sum_{i=1}^{n} \sum_{h=r_{ji}}^{a_i-1} w_{\sim kih} - d \geq s, \qquad \forall \mathbf{x}_j \in C_k \cap COR' \tag{2.50}$$

$$\sum_{i=1}^{n} \sum_{h=r_{ji}}^{a_i-1} w_{\sim kih} - \sum_{i=1}^{n} \sum_{j=1}^{r_{ji}-1} w_{kih} - d \geq s, \qquad \forall \mathbf{x}_j \in C_k^{>} \cap COR' \tag{2.51}$$

$$\sum_{i=1}^{n} \sum_{j=1}^{r_{ji}-1} w_{kih} - \sum_{i=1}^{n} \sum_{h=r_{ji}}^{a_i-1} w_{\sim kih} \leq 0 \qquad \forall \mathbf{x}_j \in C_k \cap MIS' \tag{2.52}$$

$$\sum_{i=1}^{n} \sum_{h=r_{ji}}^{a_i-1} w_{\sim kih} - \sum_{i=1}^{n} \sum_{j=1}^{r_{ji}-1} w_{kih} \leq 0, \qquad \forall \mathbf{x}_j \in C_k^{>} \cap MIS' \tag{2.53}$$

$$w_{kih} \geq z, w_{\sim kih} \geq z \tag{2.54}$$

$$\sum_{i=1}^{m} \sum_{j=1}^{a_i-1} w_{kij} = 1, \sum_{i=1}^{m} \sum_{j=1}^{a_i-1} w_{\sim kij} = 1 \tag{2.55}$$

$$d \geq 0 \tag{2.56}$$

s, t small positive constants

Constraints (2.50)–(2.51) involve only the correctly classified alternatives. In these constraints, d represents the minimum absolute difference between the global utilities of each alternative according to the two utility functions. Constraints (2.52)–(2.53) involve the alternatives misclassified after the solution of MIP (set MIS'), and it is used to ensure that they will be retained as misclassified.

After the solution of LP1, MIP, and LP2 at stage k of the hierarchical discrimination process, the "optimal" classification is achieved between the alternatives belonging in group C_k and the alternatives belonging in the groups $C_k^{>}$. The term "optimal" refers to the number of misclassifications and to the clarity of the obtained discrimination. If the current stage k is the last stage of the hierarchical discrimination process (i.e., $k = q-1$), then the model development procedure stops because all utility functions required to classify the alternatives have been estimated. Otherwise, the procedure proceeds to stage $k+1$, in order to discriminate between the alternatives belonging in group C_{k+1} and the alternatives belonging in the lower groups $C_{k+1}^{>}$. In stage $k+1$, all alternatives classified by the pair of utility functions developed at stage k into group C_k are not considered. Consequently, a new reference set A' is formed, including all alternatives that remain unclassified in a specific group (i.e., the alternatives classified in stage k in the set of groups $C_k^{>}$).

2.2.4 Model Extrapolation

The classification of a new alternative $\mathbf{x}_j \notin A'$ is performed by descending the hierarchy of Figure 2.6. Initially, the two first additive utility functions $U_1(\mathbf{g})$ and $U_{\sim 1}(\mathbf{g})$ are used to determine whether the new alternative belongs in group C_1 or not. If $U_1(\mathbf{g}_j) > U_{\sim 1}(\mathbf{g}_j)$, then $\mathbf{x}_j \in C_1$ and the procedure stops, and if $U_1(\mathbf{g}_j) < U_{\sim 1}(\mathbf{g}_j)$, then $\mathbf{x}_j \in C_1^{>}$ and the procedure proceeds with the consideration of the next pair of utility functions $U_2(\mathbf{g})$ and $U_{\sim 2}(\mathbf{g})$. If $U_2(\mathbf{g}_j) > U_{\sim 2}(\mathbf{g}_j)$, then $\mathbf{x}_j \in C_2$ and the procedure stops, and if $U_2(\mathbf{g}_j) < U_{\sim 2}(\mathbf{g}_j)$, then $\mathbf{x}_j \in C_2^{>}$ and the procedure continues in the same way until the classification of the new alternative is achieved.

2.3 Statistical and Econometric Techniques

Statistics is the oldest science involved with the analysis of given samples in order to make inferences about an unknown population. The classification problem is addressed by statistical and econometric techniques within this context. These techniques include both univariate and multivariate methods. The former involve the development and implementation of univariate statistical tests that are mainly of descriptive character. For these reasons, such techniques will not be considered in this review. The foundations of multivariate techniques can be traced back to the work of Fisher (1936) on the linear discriminant analysis (LDA). LDA has been the most extensively used methodology for developing classification models for several decades. Approximately a decade after the publication of Fisher's paper, Smith (1947) extended LDA to the more general quadratic form (quadratic discriminant analysis; QDA).

During the subsequent decades, the focus of the conducted research moved toward the development of econometric techniques. The most well-known methods from this field include the linear probability model, logistic regression and probit models. These three methods are actually special forms of regression analysis in cases where the dependent variable is discrete. The linear probability model is only suitable for two-group classification problems, whereas both logit and probit models are applicable to multi-group problems, too. The latter two methodologies have several significant advantages over discriminant analysis. This has been one of the main reasons for their extensive use.

Despite the criticism on the use of these traditional statistical and econometric approaches, they still remain quite popular both as research tools as well as for practical purposes. This popularity is supported by the existence of a plethora of statistical and econometric software, which contribute to the easy use of these approaches. Furthermore, statistical and econometric techniques are quite often considered in comparative studies investigating the performance of new classification techniques being developed. In this regard, statistical and econometric techniques often serve as a reference point (benchmark) in conducting such comparisons. It is also important to note that under specific data conditions, statistical techniques yield the optimal classification rule.

2.3.1 Discriminant Analysis

Discriminant analysis has been the first multivariate statistical classification method used for decades by researchers and practitioners in developing classification models. In its linear form it was developed by Fisher (1936). Given a training sample consisting of m alternatives whose classification is a priori known, the objective of the method is to develop a set of discriminant functions maximizing the ratio of among-groups to within-groups variance. In the general case where the classification involves q groups, $q-1$ linear functions of the following form are developed:

$$Z_{kl} = a_{kl} + b_{kl1}g_1 + b_{kl2}g_2 + \cdots + b_{kln}g_n$$

where $g_1, g_2, \ldots, g_n$ are the attributes describing the alternatives $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$, a_{kl} is a constant term, and $b_{kl1}, b_{kl2}, \ldots, b_{kln}$ are the attributes' coefficients in the discriminant function. The indices k and l refer to a pair of groups C_k and C_l.

The estimation of the model's parameters involves the estimation of the constant terms a_{kl} and the vectors $\mathbf{b}_{kl} = (b_{kl1}, b_{kl2}, \ldots, b_{kln})$. The estimation procedure is based on two major assumptions: (a) the data follow the multivariate normal distribution, and (b) the variance-covariance matrices for each group are equal. Given these assumptions, the estimation of the constant terms and the attributes' coefficients is performed as follows:

$$\begin{aligned} \mathbf{b}_{kl} &= \mathbf{S}^{-1}(\overline{\mathbf{x}}_k - \overline{\mathbf{x}}_l) \\ a_{kl} &= -(\overline{\mathbf{x}}_k + \overline{\mathbf{x}}_l)'\mathbf{b}_{kl}/2 \end{aligned}$$

where:

- $\overline{\mathbf{x}}_k$ is a $n \times 1$ vector consisting of the attributes' mean values for group C_k,
- $\mathbf{S}$ is the within-groups variance-covariance matrix, defined as follows:

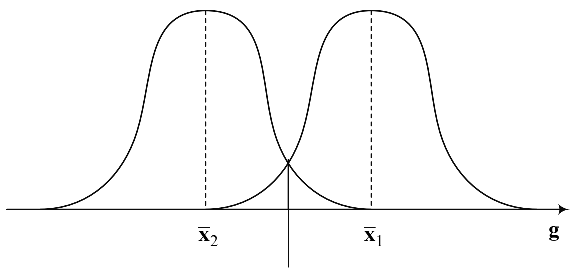

Fig. 2.8 The classification rule in linear discriminant analysis (Source: Altman et al., 1981)

$$\mathbf{S} = \frac{\sum_{k=1}^{q} \sum_{\forall \mathbf{x}_j \in C_k} (\mathbf{x}_j - \overline{\mathbf{x}}_k)(\mathbf{x}_j - \overline{\mathbf{x}}_k)'}{m-q}$$

Once the parameters (coefficients and constant term) of the discriminant functions are estimated, the classification of an alternative $\mathbf{x}_j$ is decided on the basis of its discriminant score $Z_{kl}(\mathbf{x}_j)$ assigned to the alternative by each discriminant function Z_{kl}. In particular, $\mathbf{x}_j$ is classified into group C_k if for all other groups C_l the following rule holds:

$$Z_{kl}(\mathbf{x}_j) \geq \ln \frac{K(k|l)\pi_l}{K(l|k)\pi_k}$$

In the above rule $K(k|l)$ denotes the misclassification cost corresponding to an incorrect decision to classify an alternative into group C_k while actually belong into group C_l and π_k denotes the a priori probability that an alternative belongs into group C_k. Figure 2.8 gives a graphical representation of the above linear classification rule in the two-group case, assuming that all misclassification costs and a priori probabilities are equal.

In the case where the group variance-covariance matrices are not equal, then QDA is used instead of LDA. The general form of the quadratic discriminant function developed through QDA for each pair of groups C_k and C_l is the following:

$$Z_{kl} = a_{kl} + \sum_{i=1}^{n} b_{kli} g_i + \sum_{i=1}^{n} \sum_{h=1}^{n} c_{klih} g_i g_h$$

The estimation of the coefficients and the constant term is performed as follows:

$$\begin{aligned} \mathbf{b}_{kl} &= -2(\overline{\mathbf{x}}'_k \mathbf{S}_k^{-1} - \overline{\mathbf{x}}'_l \mathbf{S}_l^{-1}) \\ \mathbf{c}_{kl} &= \mathbf{S}_k^{-1} - \mathbf{S}_l^{-1} \\ a_{kl} &= \overline{\mathbf{x}}'_k \mathbf{S}_k^{-1} \overline{\mathbf{x}}_k - \overline{\mathbf{x}}'_l \mathbf{S}_l^{-1} \overline{\mathbf{x}}_l - \ln \left| \mathbf{S}_l \mathbf{S}_k^{-1} \right| \end{aligned}$$

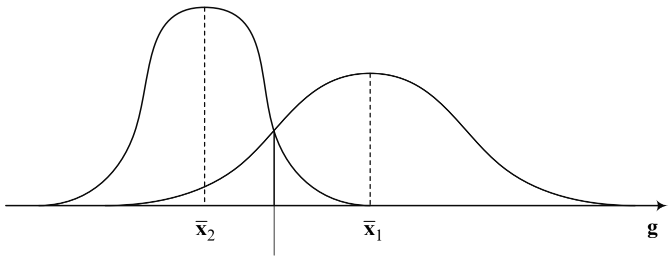

Fig. 2.9 The classification rule in quadratic discriminant analysis (Source: Altman et al., 1981)

$\mathbf{S}_k$ and $\mathbf{S}_l$ denote the within-group variance covariance matrices for groups C_k and C_l, estimated as follows:

$$\mathbf{S}_k = \frac{\sum\limits_{\forall \mathbf{x}_j \in C_k} (\mathbf{x}_j - \overline{\mathbf{x}}_k)(\mathbf{x}_j - \overline{\mathbf{x}}_k)'}{m_k - 1}$$

where m_k is the number of alternatives of the training sample that belong in group C_k.

Given the discriminant score $Z_{kl}(\mathbf{x}_j)$ of an alternative $\mathbf{x}_j$ on every discriminant function corresponding with a pair of groups C_k and C_l, the quadratic classification rule (Figure 2.9) is similar to the linear case: the alternative $\mathbf{x}_j$ is classified into group C_k if and only if for all other groups C_l the following inequality holds:

$$Z_{kl}(\mathbf{g}_j) \geq -2\ln\frac{K(k|l)\pi_l}{K(l|k)\pi_k}$$

LDA and QDA have been heavily criticized for their underlying assumptions (multivariate normality, known structure of the group variance-covariance matrices). A comprehensive discussion of the impact that these assumptions have on the obtained discriminant analysis' results is presented in the book of Altman et al. (1981).

Given that the above two major underlying assumptions are valid (multivariate normality and known structure of the group variance-covariance matrices), the use of the Bayes rule indicates that the two forms of discriminant analysis (linear and quadratic) yield the optimal classification rule (the LDA in the case of equal group variance-covariance matrices and the QDA in the opposite case). In particular, the developed classification rules are asymptotically optimal (as the training sample size increases, the statistical properties of the considered groups approximate the unknown properties of the corresponding populations). A formal proof of this finding is presented by Duda and Hart (1978), as well as by Patuwo et al. (1993).

Such restrictive statistical assumptions, however, are rarely met in practice. This fact raises a major issue regarding the real effectiveness of discriminant analysis in realistic conditions. Several studies have addressed this issue. Moore (1973), Krzanowski (1975, 1977), and Dillon and Goldstein (1978) showed that when the data include discrete variables, then the performance of discriminant analysis deteriorates especially when the attributes are significantly correlated. On the contrary, Lanchenbruch et al. (1973) and Subrahmaniam and Chinganda (1978) concluded that even in the case of non-normal data, the classification results of discriminant analysis models are quite robust, especially in the case of the QDA and for data with small degree of skewness.

2.3.2 Logit and Probit Analysis

The aforementioned problems regarding the assumptions made by discriminant analysis motivated researchers to develop more flexible methodologies. The first of such methodologies to be developed includes the linear probability model, as well as logit and probit analysis.

The linear probability model is based on a multivariate regression using as dependent variable the classification of the alternatives of the training sample. Theoretically, the result of the developed model is interpreted as the probability that an alternative belongs in one of the prespecified groups. Performing the regression, however, does not ensure that the model's result lies in the interval [0, 1], thus posing a major model interpretation problem, which makes the use of the linear probability model cumbersome, both from a theoretical and a practical perspective. For these reasons, the use of the linear probability model is rather limited and consequently it will not be further considered in this book.

Logit and probit analysis originate from the field of econometrics. Both models are based on the development of a non-linear function measuring the group-membership probability for the alternatives under consideration. The difference between the two approaches involves the form of the function that is employed. In particular, logit analysis employs the logistic function, whereas the cumulative probability density function of the normal distribution is used in probit analysis. On the basis of these functions, and assuming a two-group classification problem, the probability that an alternative $\mathbf{x}_j$ belongs in group C_2 is defined as follows[2]:

$$\text{Logit analysis: } P_j = F(a+\mathbf{g}'_j\mathbf{b})\frac{1}{1+e^{-a-\mathbf{g}'_j\mathbf{b}}} \tag{2.57}$$

$$\text{Probit analysis: } P_j = f(a+\mathbf{g}'_j\mathbf{b}) \int_{-\infty}^{a+\mathbf{g}'_j\mathbf{b}} \frac{1}{(2\pi)^{1/2}} e^{\frac{-z^2}{2}}\,dz \tag{2.58}$$

[2] If a binary 0-1 coding is used to designate each group such that $C_1 \to 0$ and $C_2 \to 1$, then equations (2.57)–(2.58) provide the probability that an alternative belongs in group C_2. If the binary coding is applied in the opposite way (i.e., $C_1 \to 1$ and $C_2 \to 0$), then equations (2.57)–(2.58) provide the probability that an alternative belongs in group C_1.

Table 2.1 The ordered logit and probit models

Ordered logit model	$P_{1j} = F\left(a_1 + \mathbf{g'}_j\mathbf{b}\right)$
	$P_{2j} = F\left(a_2 + \mathbf{g'}_j\mathbf{b}\right) - F\left(a_1 + \mathbf{g'}_j\mathbf{b}\right)$
	....
	$P_{kj} = 1 - \left(P_{1j} + P_{2j} + \cdots + P_{k-1,j}\right)$
Ordered probit model	$P_{1j} = \int_{-\infty}^{a_1+\mathbf{g'}_j\mathbf{b}} f(z)dz$
	$P_{2j} = \int_{a_1+\mathbf{g'}_j\mathbf{b}}^{a_2+\mathbf{g'}_j\mathbf{b}} f(z)dz$
	....
	$P_{kj} = \int_{a_{k-1}+\mathbf{g'}_j\mathbf{b}}^{+\infty} f(z)dz$

The estimation of the constant term a and the vector $\mathbf{b}$ is performed using maximum likelihood techniques. In particular, the parameters' estimation process involves the maximization of the following likelihood function:

$$\ln L = \sum_{\forall \mathbf{x}_j \in C_2} \ln(P_j) + \sum_{\forall \mathbf{x}_j \in C_1} \ln(1 - P_j)$$

The maximization of this function is a nonlinear optimization problem. Altman et al. (1981) report that if there exists a linear combination of the attributes $g_1, g_2, \ldots, g_n$ that accurately discriminates the prespecified groups, then the optimization process will not converge to an optimal solution.

Once the parameters' estimation process is completed, equations (2.57) and (2.58) are used to estimate the group-membership probabilities for all the alternatives under consideration. The classification decision is taken on the basis of these probabilities. For instance, in a two-group classification problem, one can impose a classification rule of the following form: "assign an alternative to group C_2 if $P_j \geq 0.5$; otherwise assign the alternative into group C_1." Alternate probability cutoff points, other than 0.5, can also be specified through trial and error processes.

In the case of multigroup classification problems, logit and probit analysis can be used in two forms: as multinomial or ordered logit/probit models. The difference among multinomial and ordered models is that the former assume a nominal definition of the groups, whereas the latter assume an ordinal definition. In this respect, ordered models are more suitable for addressing sorting problems, and traditional discrimination/classification problems are addressed through multinomial models.

The ordered models require the estimation of a vector of attributes' coefficients $\mathbf{b}$ and a vector of constant terms $\mathbf{a}$. These parameters are used to specify the probability P_{kj} that an alternative $\mathbf{x}_j$ belongs in group C_k, in the way presented in Table 2.1, where $f(z)$ is the standard normal density function.

The constant terms are defined such that $a_{k-1} > a_{k-2} > \cdots > a_2 > 0$ $(a_1 = 0)$. The parameters' estimation process is performed similar to the two-group case using maximum likelihood techniques.

The multinomial models require the estimation of a set of coefficient vectors $\mathbf{b}_k$ and a constant term a_k corresponding with each group C_k $(k = 1, 2, \ldots, q)$. On the basis of these parameters, the multinomial logit model estimates the probability P_{kj} that an alternative $\mathbf{x}_j$ belongs in group C_k as follows:

$$P_{kj} = \frac{e^{\mathbf{g}'_j \mathbf{b}_k + a_k}}{\sum_{l=1}^{q} e^{\mathbf{g}'_j \mathbf{b}_k + a_l}}$$

For normalization purposes, $\mathbf{b}_1$ and a_1 are set such that $\mathbf{b}_1 = \mathbf{0}$ and $a_1 = 0$, whereas all other $\mathbf{b}_k$ and a_k $(k = 2, \ldots, q)$ are estimated through maximum likelihood techniques.

Between the logit and probit models, the former is usually preferred. This is mainly because the development of logit models requires less computational effort. Furthermore, there are not strong theoretical and practical results to support a comparative advantage of probit models in terms of their classification accuracy.

During the past three decades, both logit and probit analyses have been extensively used by researchers in a wide range of fields as efficient alternatives to discriminant analysis. However, despite the theoretical advantages of these approaches over LDA and QDA (logit and probit analyses do not pose assumptions on the statistical distribution of the data or the structure of the group variance-covariance matrices), comparative studies made have not clearly shown that these techniques outperform discriminant analysis (linear or quadratic) in terms of their classification performance (Krzanowski, 1975; Press and Wilson, 1978).

2.4 Non-parametric Techniques

In practice, the statistical properties of the data are rarely known, because the underlying population is difficult to be fully specified. This poses problems on the use of statistical techniques and motivated researchers toward the development of non-parametric methods. Such approaches have no underlying statistical assumptions and consequently it is expected that they are flexible enough to adjust to the characteristics of the data under consideration. In the subsequent sections, the most important of these techniques are described.

2.4.1 Neural Networks

Neural networks, often referred to as artificial neural networks, have been developed by artificial intelligence researchers as an innovative modeling methodology of complex problems. The foundations of the neural networks paradigm lie on the emulation of the operation of the human brain. The human brain consists of a huge number of neurons organized in a highly complex network. Each neuron is an individual processing unit. A neuron receives an input signal (stimulus from body sensors or output signal from other neurons), which after a processing phase produces an output signal that is transferred to other neurons for further processing. The result of the overall process is the action or decision taken in accordance with the initial stimulus.

This complex biological operation constitutes the basis for the development of neural network models. Every neural network is a network of parallel processing units (neurons) organized into layers. A typical structure of a neural network (Figure 2.10) includes the following structural elements:

1. An input layer consisting of a set of nodes (processing units-neurons) one for each input to the network.
2. An output layer consisting of one or more nodes depending on the form of the desired output of the network. In classification problems, the number of nodes of the output layer is determined depending on the number of groups. For instance, for a two-group classification problem the output layer may include only one node taking two values: 1 for group C_1 and 0 for group C_2 (these are arbitrary chosen values and any other pair is possible). In the general case where there are q groups, the number of output nodes is set equal to the number of groups.
3. A series of intermediate layers referred to as hidden layers. The nodes of each hidden layer are fully connected with the nodes of the subsequent and the proceeding layer. Furthermore, it is also possible to consider more complicated structures where all layers are fully connected to each other. Such general network structures are known as fully connected neural networks. The network presented in Figure 2.10 is an example of such structure. There is no general rule to define the number of hidden layers. This is, usually, performed through trial and error processes. Recently, however, a significant part of the research has been devoted to the development of self-organizing neural network models, that is neural networks that adjust their structure to best match the given data conditions. Research made on the use of neural networks for classification purposes showed that, in many cases, a single hidden layer is adequate (Patuwo et al., 1993; Subramanian et al., 1993).

Each connection between two nodes of the network is assigned a weight representing the strength of the connection. The determination of these weights (training of the network) is accomplished through optimization techniques. The objective of the optimization process is to minimize the differences between the recommendations of the network and the actual classification of the alternatives belonging in the training sample.

The most widely used network training methodology is the back propagation approach (Rumerlhart et al., 1986). Recently, advanced nonlinear optimization techniques have also contributed to obtaining globally optimum estimations of the network's connection weights (Hung and Denton, 1993).

On the basis of the connections' weights, the input to each node is determined as the weighted average of the outputs of all other nodes with which there are established connections. In the general case of a fully connected neural network (cf. Figure 2.10) the input in_{ir} to node i of the hidden layer r is defined as follows:

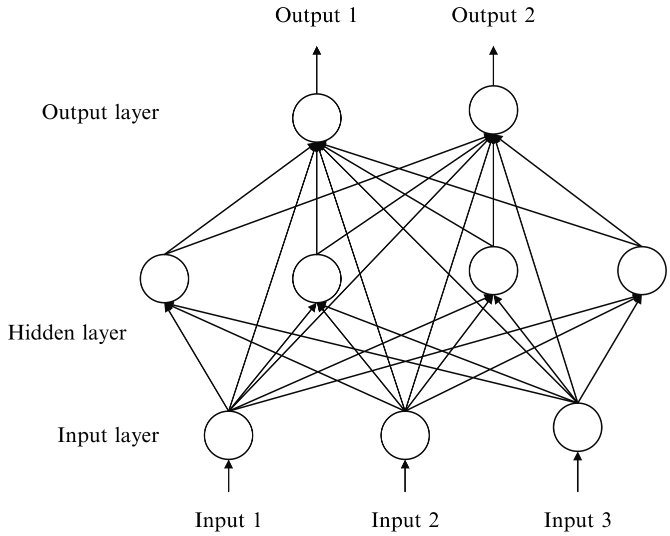

Fig. 2.10 A general architecture of a neural network

$$in_{ir} = \sum_{j=0}^{r-1} \sum_{k=1}^{n_j} w_{ik}^{j} o_{kj} + \phi_{ir}$$

where:

n_j the number of nodes at the hidden layer j,

w_{ik} the weight of the connection between node i of layer r and node k of layer j,

o_{kj} the output of node k at layer j,

ϕ_{ir} an error term.

The output of each node is specified through a transformation function. The most common form of this function is the logistic function:

$$o_{ir} = \frac{1}{1 + e^{-\frac{in_{ir}}{T}}}$$

where T is a user-defined constant.

The major advantage of neural networks is their parallel processing ability as well as their ability to represent highly complex, nonlinear systems. Theoretically, this enables the approximation of any real function with infinite accuracy (Kosko, 1992). These advantages led to the widespread application of neural networks in many research fields. On the other hand, the criticism of the use of neural networks is focused on two points:

1. The increased computational effort required for training the network (specification of connections' weights).

2. The inability to provide explanations of the network's results. This is a significant shortcoming, mainly from a decision support perspective, as in a decision-making context, the justification of the final decision is often a crucial point.

Except for the above two problems, research studies investigating the classification performance of neural networks as opposed to statistical and econometric techniques have led to conflicting results. Subramanian et al. (1993) compared neural networks with LDA and QDA through a simulation experiment using data conditions that were in accordance with the assumptions of the two statistical techniques. Their results show that neural networks can be a promising approach, especially in cases of complex classification problems involving more than two groups and a large set of attributes. On the other hand, LDA and QDA performed better when the sample size was increased.

A similar experimental study by Patuwo et al. (1993) leads to the conclusion that there are many cases where statistical techniques outperform neural networks. In particular, the authors compared neural networks with LDA and QDA, considering both the case where the data conditions are in line with the assumptions of these statistical techniques, as well as the opposite case. According to the obtained results, when the data are multivariate normal with equal group variance-covariance matrices, then LDA outperforms neural networks. Similarly in the case of multivariate normality with unequal variance-covariance matrices, QDA outperformed neural networks. Even in the case of non-normal data, the results of the analysis did not show any clear superiority of neural networks, at least compared with QDA.

The experimental analysis of Archer and Wang (1993) is also worth mentioning. The authors discussed the way that neural networks can be used to address sorting problems and compared their approach with LDA. The results of this comparison show a higher classification performance for the neural networks approach, especially when there is a significant degree of group overlap.

2.4.2 Rule Induction and Decision Trees

During the past two decades, machine learning evolved as a major discipline within the field of artificial intelligence. Its objective is to describe and analyze the computational procedures required to extract and organize knowledge from the existing experience. Within the different learning paradigms (Kodratoff and Michalski, 1990), inductive learning through examples is the one most widely used.

In contrast with the classification techniques described in the previous sections, inductive learning introduces a completely different approach in modeling the classification problem. In particular, inductive learning approaches organize the extracted knowledge in a set of decision rules of the following general form:

IF *elementary conditions* THEN *conclusion*

The first part of such rules examines the necessary and sufficient conditions required for the conclusion part to be valid. The elementary conditions are connected

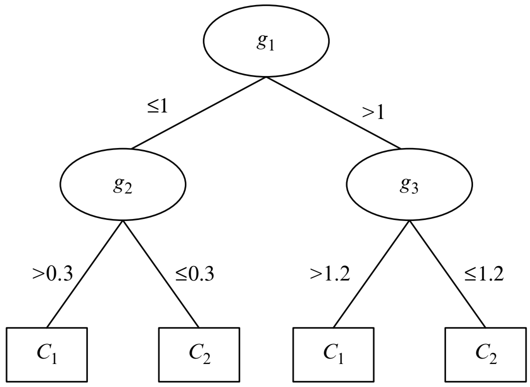

Fig. 2.11 A sample classification decision tree developed using the C4.5 algorithm

using the AND operator. The conclusion consists of a recommendation on the classification of the alternatives satisfying the conditions part of the rule.

One of the most widely used techniques developed on the basis of the inductive learning paradigm is the C4.5 algorithm (Quinlan, 1993). The decision rules developed through the C4.5 algorithm are organized in the form of a decision tree such as the one presented in Figure 2.11. Every node of the tree considers an attribute, and the branches correspond with elementary conditions defined on the basis of the node attributes. Finally, the leaves designate the group to which an alternative is assigned, given that it satisfies the branches' conditions.

The development of the classification tree is performed through an iterative process. Every stage of this process consists of three individual steps:

1. Evaluation of the discriminating power of the attributes in classifying the alternatives of the training sample.
2. Selection of the attribute having the highest discriminating power.
3. Definition of subsets of alternatives on the basis of their performances on the selected attribute.

This procedure is repeated for every subset of alternatives formed in the third step, until all alternatives of the training sample are correctly classified. The evaluation of the attributes' discriminating power in the first step of the above process is performed on the basis of the amount of new information introduced by each attribute in the classification of the alternatives.

The entropy of the classification introduced by each attribute is used as the appropriate information measure. In particular, assuming that each attribute introduces a partitioning of the training sample into t subsets $D_1, D_2, \ldots, D_t$, each consisting of v_t alternatives, then the entropy of this partitioning is defined as follows:

$$I(D) = -\sum_{h=1}^{t} \frac{v_h}{m} \sum_{k=1}^{q} p(D_h/C_k) \log_2[p(D_h/C_k)]$$

where, $p(D_h/C_k)$ denotes the number of alternatives of set D_h that belong in group C_k. The attribute with the minimum entropy is selected as the one with the highest discriminating power. This attribute adds the highest amount of new information in the classification of the alternatives.

The above procedure may lead to a highly specialized classification tree with nodes covering only one alternative. This is the result of overfitting the tree to the given data of the training sample, a phenomenon that is often related to poor generalizing performance. C4.5 addresses this problem through the implementation of a pruning phase, so that the decision tree's size is reduced, in order to improve its expected generalizing performance. The development and implementation of pruning methodologies is a significant research topic in the machine learning community. Some characteristic examples of pruning techniques are the ones presented by Breiman et al. (1984), Gelfand et al. (1991), and Quinlan (1993).

The general aspects of the paradigm used in C4.5 are common to other machine learning algorithms. Some well-known examples of such algorithms include CN2 (Clark and Niblett, 1989), the AQ family of algorithms (Michalski, 1969), and the recursive partitioning algorithm (Breiman et al., 1984).

The main advantages of machine learning classification algorithms involve the following capabilities:

1. Handling of qualitative attributes.
2. Flexibility in handling missing information.
3. Exploitation of large data sets for model development purposes through computationally efficient procedures.
4. Development of easily understandable classification models.

2.4.3 Fuzzy Set Theory

Decision making is often based on fuzzy, ambiguous and vague judgments. The daily use of verbal expressions such as "almost," "usually," "often," etc., are simple yet typical examples of this remark. The fuzzy nature of these simple verbal statements is indicative of the fuzziness encountered in the decision-making process. The fuzzy set theory developed by Zadeh (1965) provides the necessary modeling tools for the representation of uncertainty and fuzziness in complex real-world situations.

The core of this innovative approach is the fuzzy set concept. A fuzzy set is a set with no crisp boundaries. In the case of a traditional crisp set a proposition of the form "alternative $\mathbf{x}$ belongs to the set A" is either true or false; for a fuzzy set, however, it can be partly true or false. Within the context of the fuzzy set theory, the modeling of such fuzzy judgments is performed through the definition of membership functions. A membership function defines the membership degree that an object (alternative) belongs in a fuzzy set. The membership degree ranges in the interval [0, 1]. In the aforementioned example, a membership degree equal to 1 indicates that the proposition "alternative $\mathbf{x}$ belongs to the set A" is true. Similarly, if the membership degree is 0, then it is concluded that the proposition is false. Any other

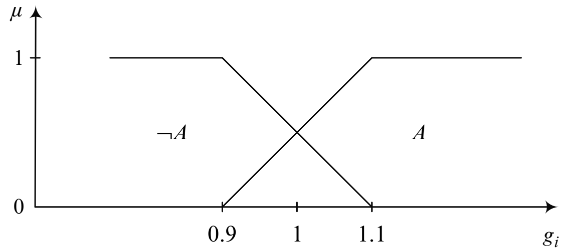

Fig. 2.12 An example of a membership function

value for the membership degree between 0 and 1 indicates that the proposition is partly true.

Figure 2.12 presents an example of a typical form for the membership function for the proposition "according to attribute g_i, alternative $\mathbf{x}$ belongs to the set A." The membership function corresponding with the negation of this proposition is also presented (the negation defines the complement set of A, denoted as $\neg A$).

In order to derive an overall conclusion regarding the membership of an alternative into a fuzzy set based on the consideration of all attributes, one must aggregate the partial membership degrees for each individual attribute. This aggregation is based on common operators such as "AND" and "OR" operators. The former corresponds with the union operation, whereas the latter indicates a intersection operation. A combination of these two operators is also possible.

In the case of classification problems, each group can be considered as a fuzzy set. Similar to the machine learning paradigm, classification models developed through approaches that implement the fuzzy set theory have the form of decision rules. The general form of a fuzzy rule used for classification purposes is the following:

$$\text{IF } (g_{j1} \text{ is } A_{1a}) \wedge (g_{j2} \text{ is } A_{2b}) \wedge \ldots \wedge (g_{jn} \text{ is } A_{nc}) \text{ THEN } \mathbf{x}_j \in C_k$$

where each $A_{i(\cdot)}$ corresponds with a fuzzy set defined on the scale of attribute g_i. The strength of each individual condition is defined by the membership degree of the corresponding proposition "according to attribute g_j alternative $\mathbf{x}_j$ belongs to the set $A_{i(\cdot)}$." The rules of the above general form are usually associated with a certainty coefficient indicating the certainty about the validity of the conclusion part.

Procedures for the development of fuzzy rules in classification problems have been proposed by several researchers. Some indicative studies on this field are the ones of Ishibuchi et al. (1992, 1993), Inuiguchi et al. (2000), Bastian (2000), and Oh and Pedrycz (2000).

Despite the existing debate on the relation between the fuzzy set theory and the traditional probability theory, fuzzy sets have been extensively used to address a variety of real-world problems from several fields. Furthermore, several researchers have exploited the underlying concepts of the fuzzy set theory in conjunction with other disciplines such as neural networks (neurofuzzy systems; Von Altrock, 1996), expert systems (fuzzy rule-based expert systems; Langholz et al., 1996), mathematical programming (fuzzy mathematical programming; Zimmermann, 1978), and MCDA (Yager, 1977; Dubois and Prade, 1979; Siskos, 1982; Siskos et al., 1984; Fodor and Roubens, 1994; Grabisch, 1995, 1996; Lootsma, 1997).

2.4.4 Rough Sets

Pawlak (1982) introduced the rough set theory as a tool to describe dependencies between attributes, to evaluate the significance of attributes, and to deal with inconsistent data. As an approach to handle imperfect data (uncertainty and vagueness), it complements other theories that deal with data uncertainty, such as probability theory, evidence theory, fuzzy set theory, etc. Generally, the rough set approach is a very useful tool in the study of classification problems. Recently, however, there have been several advances in this field to allow the application of the rough set theory to choice and ranking problems as well (Greco et al., 1997).

The rough set philosophy is founded on the assumption that with every alternative, some information (data, knowledge) is associated. This information involves two types of attributes: condition and decision attributes. Condition attributes are those used to describe the characteristics of the objects. For instance, the set of condition attributes describing a country can be a set of economic, political, and social indicators. The decision attributes define a partition of the objects into groups according to the condition attributes.

On the basis of these two types of attributes, an information table $S = \langle U, Q, V, f \rangle$ is formed, as follows:

- U is a finite set of m alternatives (objects).
- Q is a finite set of n attributes.
- V is the intersection of the domains of all attributes (the domain of each attribute g_i is denoted by V_i). The traditional rough set theory assumes that the domain of each attribute is a discrete set. In this context, every quantitative real-valued attribute needs to be discretized[3] using discretization algorithms such as the ones proposed by Fayyad and Irani (1992), Chmielewski and Grzymala-Busse (1996), and Zighed et al. (1998). Recently, however, the traditional rough set approach has been extended so that no discritezation is required for quantitative attributes. Typical examples of the new direction are the DOMLEM algorithm (Greco et al., 1999a) and the MODLEM algorithm (Grzymala-Busse and Stefanowski, 2001).
- $f : U \times Q \to V$ is a total function such that $f(\mathbf{x}_j, g_i) \in V_i$ for every $g_i \in Q$, $\mathbf{x}_j \in U$, called information function (Pawlak, 1991; Pawlak and Slowinski, 1994).

Simply stated, the information table is an $m \times n$ matrix, with rows corresponding with the alternatives and columns corresponding with the attributes.

Given an information table, the basis of the traditional rough set theory is the indiscernibility between the alternatives. Two alternatives $\mathbf{x}_j$ and $\mathbf{x}_l$ are considered

[3] Discretization involves the partitioning of an attribute's domain $[a,b]$ into h subintervals $[v_1, v_2)$, $[v_2, v_3)$, ..., $[v_{h-1}, v_h]$, where $v_1 = a$ and $v_h = b$.

to be indiscernible, if and only if they are characterized by the same information, i.e., $f(\mathbf{x}_j, g_i) = f(\mathbf{x}_l, g_i)$ for every $g_i \in P \subseteq Q$. In this way, every $P \subseteq Q$ leads to the development of a binary relation on the set of alternatives. This relation is called $P-$indiscernibility relation, denoted by I_P. I_P is an equivalence relation for any P.

Every set of indiscernible alternatives is called elementary set and it constitutes a basic granule of knowledge. Equivalence classes of the relation I_P are called $P-$elementary sets in S, and $I_P(\mathbf{x}_j)$ denotes the $P-$elementary set containing alternative $\mathbf{x}_j \in U$.

Any set of objects being a union of some elementary sets is referred to as crisp (precise) otherwise it is considered to be rough (imprecise, vague). Consequently, each rough set has a boundary line consisting of cases (objects) that cannot be classified with certainty as members of the set or of its complement. Therefore, a pair of crisp sets, called the lower and the upper approximation can represent a rough set. The lower approximation consists of all objects that certainly belong to the set and the upper approximation contains objects that possibly belong to the set. The difference between the upper and the lower approximation defines the doubtful region, which includes all objects that cannot be certainly classified into the set. On the basis of the lower and upper approximations of a rough set, the accuracy of its approximation can be calculated as the ratio of the cardinality of its lower approximation to the cardinality of its upper approximation.

Assuming that $P \subseteq Q$ and $Y \subseteq U$, then the $P-$lower approximation, the $P-$upper approximation, and the $P-$doubtful region of Y ($\underline{P}Y$, $\overline{P}Y$, and $BN_P(Y)$, respectively) are formally defined as follows:

$$\underline{P}Y = \{\mathbf{x}_j \in Y : I_P(\mathbf{x}_j) \in Y\} \tag{2.59}$$

$$\overline{P}Y = \bigcup_{\mathbf{x}_j \in Y} I_p(\mathbf{x}_j) \tag{2.60}$$

$$BN_P(Y) = \overline{P}Y - \underline{P}Y \tag{2.61}$$

On the basis of these approximations, it is possible to estimate the accuracy of the approximation of the rough set Y, denoted by $\alpha_P(Y)$. The accuracy of the approximation is defined as the ratio of the number of alternatives belonging in the lower approximation to the number of alternatives of the upper approximation:

$$\alpha_P(Y) = \frac{|\underline{P}Y|}{|\overline{P}Y|}$$

Within the context of a classification problem, each group C_k is considered as a rough set $_k$. The overall quality of the approximation of the classification by a set of attributes P is defined as follows:

$$\gamma_P(Y) = \frac{\sum_{k=1}^{q} |\underline{P}Y_k|}{m}$$

Having defined the quality of the approximation, the first major capability that the rough set theory provides is to reduce the available information, so as to retain only the information that is absolutely necessary for the description and classification of the alternatives. This is achieved by discovering subsets R of the complete set of attributes P, which can provide the same quality of classification as the whole attributes' set, i.e., $\gamma_P(Y) = \gamma_R(Y)$. Such subsets of attributes are called reducts and are denoted by $\mathrm{RED}_Y(P)$. Generally, the reducts are more than one. In such a case, the intersection of all reducts is called the core, i.e., $\mathrm{CORE}_Y(P) = \cap \mathrm{RED}_Y(P)$. The core is the collection of the most relevant attributes, which cannot be excluded from the analysis without reducing the quality of the obtained description (classification). The decision maker can examine all obtained reducts and proceed to the further analysis of the considered problem according to the reduct that best describes reality. Heuristic procedures can also be used to identify an appropriate reduct (Slowinski and Zopounidis, 1995).

The subsequent steps of the analysis involve the development of a set of rules for the classification of the alternatives into the groups where they actually belong. The rules developed through the rough set approach have the following form:

IF *conjunction of elementary conditions*
THEN *disjunction of elementary decisions*

The procedures used to construct a set of decision rules employ the machine learning paradigm. Such procedures developed within the context of the rough set theory have been presented by Grzymala-Busse (1992), Slowinski and Stefanowski (1992), Skowron (1993), Ziarko et al. (1993), Stefanowski and Vanderpooten (1994), Mienko et al. (1996), and Grzymala-Busse and Stefanowski (2001). Generally, rule induction techniques follow one of the following strategies:

1. Development of a minimal set of rules covering all alternatives of the training sample (information table).
2. Development of an extensive set of rules consisting of all possible decision rules.
3. Development of a set of strong rules, even partly discriminant,[4] which do not necessarily cover all alternatives of the training sample.

Irrespective of the rule induction approach employed, a decision rule developed on the basis of the rough set approach has some interesting properties and features. In particular, if all alternatives that satisfy the condition part belong in the group indicated by the conclusion of the rule, then the rule is called consistent. In the case where the condition part considers only a single group, then the rule is called exact, otherwise the rule is called approximate. The conclusion part of approximate

[4] Rules covering only alternatives that belong to the group indicated by the conclusion of the rule (positive examples) are called discriminant rules. On the contrary, rules that cover both positive and negative examples (alternatives not belonging in the group indicated by the rule) are called partly discriminant rules. Each partly discriminant rule is associated with a coefficient measuring the consistency of the rule. This coefficient is called level of discrimination and is defined as the ratio of positive to negative examples covered by the rule.

rules involves a disjunction of at least two groups ($\mathbf{x}_j \in C_k \vee \mathbf{x}_j \in C_h \vee \ldots$). Approximate rules are developed when the training sample (information table) includes indiscernible alternatives belonging in different groups. Each rule is associated with a strength measure, indicating the number of alternatives covered by the rule. For approximate rules, their strength is estimated for each individual group considered in their conclusion part. Stronger rules consider a limited number of elementary conditions; thus, they are more general.

Once the rule induction process is completed, the developed rules can be easily used to decide upon the classification of any new alternative not considered during model development. This is performed by matching the conditions part of each rule to the characteristics of the alternative, in order to identify a rule that covers the alternative. This matching process may lead to one of the following four situations (Slowinski and Stefanowski, 1994):

1. The alternative is covered only by one exact rule.
2. The alternative is covered by more than one exact rule, all indicating the same classification.
3. The alternative is covered by one approximate rule or by more than one exact rule indicating different classifications.
4. The alternative is not covered by any rule.

The classification decision in situations (1) and (2) is straightforward. In situation (3), the developed rule set leads to conflicting decisions regarding the classification of the alternative. To overcome this problem, one can consider the strength of the rules that cover the alternative (for approximate rule, the strength for each individual group of the condition part must be considered). The stronger rule can be used to take the final classification decision. This approach is employed in the LERS classification system developed by Grzymala-Busse (1992).

Situation (4) is the most difficult one, because using the developed rule set one has no evidence as to the classification of the alternative. The LERS system tackles this problem through the identification of rules that partly cover the characteristics of the alternative under consideration.[5] The strength of these rules as well as the number of elementary conditions satisfied by the alternative are considered in making the decision. An alternative approach proposed by Slowinski (1993) involves the identification of a rule that best matches the characteristics of the alternative under consideration. This is based on the construction of a valued closeness relation measuring the similarity between each rule and the alternative. The construction of this relation is performed in two stages. The first stage involves the identification of the attributes that are in accordance to the affirmation "the alternative $\mathbf{x}_j$ is close to rule r." The strength of this affirmation is measured on a numerical scale between 0 and 1. The second stage involves the identification of the characteristics that are in discordance with the above affirmation. The strength of concordance and discordance tests are combined to estimate an overall index representing the similarity of a rule to the characteristics of the alternative.

[5] Partly covering involves the case where the alternative satisfies only some of the elementary conditions of a rule.

Closing this brief discussion of the rough set approach, it is important to note the recent advances made in this field toward the use of the rough set approach as a methodology of preference modeling in multicriteria decision problems (Greco et al., 1999a, 2000a). The main novelty of the recently developed rough set approach concerns the possibility of handling criteria, i.e., attributes with preference ordered domains, and preference ordered groups in the analysis of sorting examples and the induction of decision rules. The rough approximations of decision groups involve dominance relation, instead of indiscernibility relation considered in the basic rough set approach. They are built of reference alternatives given in the sorting example (training sample). Decision rules derived from these approximations constitute a preference model. Each "if ... then ..." decision rule is composed of (a) a condition part specifying a partial profile on a subset of criteria to which an alternative is compared using the dominance relation, and (b) a decision part suggesting an assignment of the alternative to "at least" or "at most" a given class.

The decision rule preference model has also been considered in terms of conjoint measurement (Greco et al., 2001). A representation theorem for multicriteria sorting proved by Greco et al. states an equivalence of simple cancellation property, a general discriminant (sorting) function, and a specific outranking relation, on the one hand, and the decision rule model on the other hand. It is also shown that the decision rule model resulting from the dominance-based rough set approach has an advantage over the usual functional and relational models because it permits handling inconsistent sorting examples. The inconsistency in sorting examples is not unusual due to instability of preference, incomplete determination of criteria, and hesitation of the decision maker.

It is also worth noting that the dominance-based rough set approach is able to deal with sorting problems involving both criteria and regular attributes whose domains are not preference ordered (Greco et al., 2002) and missing values in the evaluation of reference alternatives (Greco et al., 1999b; Greco et al., 2000b). It also handles ordinal criteria in a more general way than the Sugeno integral, as it has been proved in Greco et al. (2001).

2.5 Miscellaneous Techniques

Apart from the techniques that have been described above, other multi-criteria optimization methodologies could also be used for the evaluation of country risk. Based on Olson and Shi (2005), financial management problems can be data mined using large real-life data sets. Moreover, they support that in the financial business, practitioners have applied a number of data-mining techniques to support credit card portfolio management and further financial management. These techniques include the Behavior Score developed by Fair Isaac Corporation (FICO), Credit Bureau Scores, First Data Resource (FDR)'s Proprietary Bankruptcy Score and Set Enumeration (SE) decision tree.

Kou et al. (2003) promote a multiple criteria linear programming (MCLP) approach to data mining based on linear discriminant analysis. They describe the connections between MCLP and data mining, including several general models of MCLP approaches. Similarly, Kou et al. (2004) propose a classification model by using multiple criteria linear programming to discover behavior patterns of credit card holders. As continuation of this research, He et al. (2004) propose a heuristic classification method by using the fuzzy linear programming to discover the bankruptcy patterns of credit card holders.

Taking into account that credit risk and bankruptcy risk constitute part of country financial risk, the aforementioned methodologies could also be used for the evaluation of country risk.

Chapter 3
Applications

This chapter includes the major characteristics of few studies that have been already done and contribute significantly to country risk analysis.

3.1 The Study of Zopounidis and Doumpos (1997)

The MCDA methodologies that have already been applied in country risk assessment studied the problem either from the ranking point of view or the portfolio construction point of view. Their aim was to develop multicriteria decision models in order to rank a set of countries from the less to the more risky ones or to develop models that could be used to construct a portfolio of countries that maximizes the return of an investment and minimizes the associated risk.

In this case, the UTASTAR and the UTADIS, I, II, and III methods are applied in the assessment of country risk, in order to develop country risk models for the ranking and sorting of a set of 66 countries according to their economic performance.

3.1.1 Data Set Description

This application involves the assessment of the country risk of 66 countries from different geographical regions all around the world. More specifically, the sample data includes 18 European countries, 16 countries from Asia, 15 countries from Africa, 15 countries from America, and finally two countries from Oceania. These countries were selected among the 133 countries that are included in the World Bank tables. The selection was based on the availability of the data of the countries, in order to have a complete sample of data. The period of the analysis involves the year 1994. The data of this specific year were the most recent that could be obtained during the period that this research was conducted.

K. Kosmidou et al. (eds.) *Country Risk Evaluation,* DOI: 10.1007/978-0-387-76680-5_3,

The countries are evaluated along 12 criteria, including 10 economic indicators, the political risk, as well as a development level indicator concerning the life expectancy. The data concerning the economic indicators and the life expectancy were drawn from the World Bank tables (World Bank Development Indicators of 1996), and the data regarding the political risk were drawn for the estimations of Euromoney. More specifically, the 12 evaluation criteria that are used in this case study are the following:

1. Current account balance as percentage of Gross National Product (GNP): This variable is related to the probability of default, as the current account deficit represents the amount of new financing that a country requires. Consequently, countries with large account deficits are more likely to default.
2. Exports average annual growth rate: For most countries, especially those of high-income economies, exports are the main source of foreign exchange earnings. Consequently, countries with high average annual growth rate are more capable in meeting their commitments regarding their foreign debt. The computation of this criterion was based on the exports of the countries during the period 1980–1994.
3. Imports average annual growth rate: Unlike exports, imports lead to loss of foreign exchange earnings. This criterion represents the imports' average annual growth rate during the period 1980–1994.
4. GNP per capita: GNP per capita is a very common criterion used in country risk assessment. It represents the country's level of development, indicating the flexibility of a country in reducing the consumption. Countries with a low-income economy are expected to be more inflexible in reducing the consumption, which can result in debt service difficulties and therefore default.
5. Average annual growth rate of GNP per capita: This criterion provides a dynamic measurement of the development of a country. In this case study it represents the evolution of the GNP per capita during the period 1980–1994.
6. Gross domestic investment: Gross domestic investment is strictly related to the development of a country. Domestic investments contribute directly to the GNP growth, and furthermore they are conducive to the decrease of unemployment (Calverley, 1990).
7. External debt as percentage of GNP: External debt represents the commitments of each country to its debtors. This ratio represents the size of the debt in relation to the economy's resources. Therefore, the higher the ratio, the greater the probability of a country to default.
8. Gross international reserves as percentage of GNP: Gross international reserves (international reserves excluding gold) are the main mean for servicing foreign debt. Developed countries are expected to have more international reserves available than countries of low-income economies.
9. Reserves to imports ratio: Possible fluctuations in foreign exchange receipts may result in significant debt-servicing problems for a country. On the other hand, reserves provide a protection to such fluctuations, at least for the short-term, as the larger reserves to imports, the larger is the amount of reserves that is available for the payment of the external debt.

10. Net foreign debt to exports ratio: As already mentioned, exports constitute the main source of foreign exchange earnings for a country. On the other hand, net foreign debt, measured as the foreign debt minus reserves, represents the debt load of a country. Therefore, a high net foreign debt to exports ratio means that the country could be exposed to significant debt-servicing problems due to foreign exchange earning crises.
11. Life expectancy: Life expectancy provides an acceptable general measure of the socioeconomic development of countries. Countries of significant economic as well as social development are expected to have high life expectancy, whereas on the contrary the life expectancy of countries facing essential social and economic problems is low.
12. Political risk: The evaluation of the countries according to their political risk was drawn from Euromoney. Euromoney polls risk analysts, risk insurance brokers, and bank credit officers and asks them to give each country a score between 25 and zero. A score of 25 indicates no political risk, and zero indicates that there is high political risk. Countries are scored in comparison both with each other and with previous years.

The criteria involving the imports' average annual growth rate, the external debt as percentage of GNP, and the net foreign debt to exports have negative rates, which means that the higher the values of these criteria, the more likely it is for a country to default. On the contrary, all the other evaluation criteria have positive rates, which means that the higher the values of these criteria, the higher is the overall economic performance of a country.

The World Bank tables include also many other indicators (more or less significant) regarding the overall economic performance of each country, including detailed trade indicators, economic growth indicators, external economic indicators, and balance of payments indicators, among others. However, the evaluation of country risk in this case study had to be based on a finite, flexible, and acceptable set of evaluation criteria that could sufficiently describe the overall socioeconomic and political situation in each country. Hence, the aforementioned 12 evaluation criteria were selected upon their relevance in country risk assessment based on previous studies that have been presented by academic researchers in this field (Mumpower et al., 1987; Cosset and Roy, 1989; Oral et al., 1992; Cosset et al., 1992).

The World Bank, apart from the valuable data that it provides concerning the indicators that affect the countries' socioeconomic development, also provides a grouping of the countries based mainly on their economic performance. More specifically, the World Bank classifies the countries in four major groups:

- High-income economies (group C_1): This group includes 20 countries, mostly Western European ones, as well as the United States, Canada, Japan, Australia, New Zealand, and Israel. These countries are considered as the world's top economies, with a stable political and social environment.
- Upper-middle income economies (group C_2): Ten countries are included in this second group. These countries cannot be considered as developed ones neither from the economic nor from the sociopolitical point of view. However, they do

have some positive perspectives for future development. The countries that belong in this group include two European countries (Greece and Hungary), Southeast Asian countries such as South Korea and Malaysia, as well as countries located in Latin and South America such as Mexico, Brazil, Chile, etc.

- Lower-middle income economies (group C_3): This group includes 18 countries, located in Europe (Romania and Poland), Asia (Indonesia, Philippines, Jordan, Thailand, Turkey, etc.), Africa (Morocco, Algeria, Tunisia), and South and Latin America (Bolivia, Guatemala, Ecuador, El Salvador, Peru, etc.). These countries are facing economic as well as social and political problems, which make their future doubtful and uncertain.
- Low-income economies (group C_4): This final group consists of 18 countries facing significant problems from any aspect (economic, political, or social). Such countries include Asian countries (Nepal, Bangladesh, India, Pakistan, etc.), African countries (Kenya, Mali, Nigeria, Senegal, etc.), and Nicaragua.

This grouping of the countries was used as input to the UTADIS, I, II, and III methods, in order to develop classification country risk models representing the evaluation methodology and policy that is followed by the top officers of the World Bank.

Moreover, in order to develop a ranking country risk model through the UTASTAR method, the country risk rating of Euromoney was used. Euromoney provides country risk assessments based on nine categories of indicators that fall into three broad groups: analytical, credit, and market indicators. These indicators include the economic data of the countries, their political risk, debt indicators, credit ratings, and access to capital markets among others. Based on these indicators, a simple weighted average model is used to rank the countries according to their creditworthiness from the best to the worst ones. The country risk rating provided by Euromoney is considered as a reliable estimation, which has already been used in many previous studies of country risk assessment.

This Euromoney country risk rating was used as input to the UTASTAR method in order to develop the country risk model to rank the countries according to their creditworthiness. It is worth noting that the ranking provided by Euromoney depicts some differences compared with the grouping provided by World Bank. Some countries that the World Bank considers to be in the low-income group, such as China and India, according to Euromoney have a higher country risk rating than most of the countries that the World Bank considers to be in the lower-middle group and even than some of countries in the upper-middle income group.

3.1.2 Presentation of Results

Following the methodology that was described in Chapter 2, the UTADIS, I, II, III and UTASTAR methods were applied in the sample data of the 66 countries under consideration to develop sorting and ranking country risk models according to the grouping and the ranking provided by World Bank and Euromoney, respectively. The obtained results of the five methods are presented below.

Results of the UTADIS Method

The additive utility model developed through the UTADIS method is fully consistent with the predefined grouping of the countries according to their economic performance, which is related to the risk and the creditworthiness of a country. All countries are classified by the model in the group to which they actually belong, resulting in a classification accuracy of 100%.

Furthermore, the model also provides the competitive level between the countries of the same class. More specifically, according to the global utilities of the countries, the most creditworthy and economically developed ones are Switzerland, Norway, Belgium, The Netherlands, Denmark, and Japan. The global utilities of these countries were over 0.96. South Korea and Greece were found to be the best countries among the upper-middle income group. South Korea is located in a geographical region (East Asia) that evolved significantly the past decades, while Greece as a member of the European Union has received considerable economic support. These are the basic characteristics that distinguish these two specific countries from the other upper-middle income economies. Thailand, Costa Rica, and Peru were found to be the best in the lower-middle income group. Finally, Nicaragua, Malawi, and Cameroon were found to be countries with the higher country risk. Table 3.1 presents in detail the obtained results, as well as the original and the estimated classification of the countries ($\widehat{C}$ and C respectively).

The GNP per capita was found to be the dominant factor in the classification of the countries, with a weight of over 50% (52.14%). This is in accordance with the findings of other studies related to country risk assessment, which have also concluded in the same result (Cosset and Roy, 1989; Oral et al., 1992). The rest of the evaluation criteria have rather similar significance in the developed classification model, ranging from 1.41% for the exports' average annual growth rate to 7.97% for the net foreign debt/exports ratio. Furthermore, the significance of the GNP per capita in the classification of the countries in this case study is also confirmed by the fact that according to the data of the 66 countries under consideration, this specific criterion is able to provide an accurate classification. More specifically, all the high-income economies have a GNP per capita over \$9,320 (Portugal); the GNP per capita for the upper-middle income economies ranges between \$2,970 (Brazil) and \$8,260 (South Korea). Similarly, the GNP per capita of the lower-middle and low-income economies ranges between \$770 (Bolivia) and \$2,500 (Turkey) and \$170 (Malawi) and \$720 (Egypt), respectively.

Table 3.1 Classification results obtained through the UTADIS method

Countries	Estimated class	Utility	Actual class	Countries	Estimated class	Utility	Actual class
Switzerland	C_1	0.969	C_1	Tunisia	C_3	0.573	C_3
Norway	C_1	0.966	C_1	Poland	C_3	0.568	C_3
Belgium	C_1	0.965	C_1	Turkey	C_3	0.529	C_3
Netherlands	C_1	0.962	C_1	El Salvador	C_3	0.529	C_3
Denmark	C_1	0.962	C_1	Algeria	C_3	0.528	C_3
Japan	C_1	0.961	C_1	Ecuador	C_3	0.517	C_3
Italy	C_1	0.961	C_1	Papua-New Guinea	C_3	0.513	C_3
Australia	C_1	0.961	C_1	Jordan	C_3	0.511	C_3
Austria	C_1	0.960	C_1	Guatemala	C_3	0.510	C_3
United States	C_1	0.960	C_1	Dominican Republic	C_3	0.509	C_3
France	C_1	0.959	C_1	Morocco	C_3	0.506	C_3
United Kingdom	C_1	0.959	C_1	Romania	C_3	0.503	C_3
Sweden	C_1	0.959	C_1	Indonesia	C_3	0.495	C_3
Finland	C_1	0.958	C_1	Philippines	C_3	0.493	C_3
Canada	C_1	0.952	C_1	Bolivia	C_3	0.464	C_3
Israel	C_1	0.946	C_1	u_3		0.464	
Ireland	C_1	0.940	C_1	Egypt	C_4	0.462	C_4
New Zealand	C_1	0.938	C_1	Sri Lanka	C_4	0.462	C_4
Spain	C_1	0.937	C_1	India	C_4	0.439	C_4
Portugal	C_1	0.917	C_1	Pakistan	C_4	0.436	C_4
u_1		0.917		China	C_4	0.433	C_4
Korea, Rep.	C_2	0.915	C_2	Ghana	C_4	0.430	C_4
Greece	C_2	0.914	C_2	Senegal	C_4	0.425	C_4
Uruguay	C_2	0.810	C_2	Bangladesh	C_4	0.417	C_4
Mexico	C_2	0.778	C_2	Kenya	C_4	0.417	C_4
Hungary	C_2	0.748	C_2	Nepal	C_4	0.413	C_4
Chile	C_2	0.731	C_2	Ivory Coast	C_4	0.399	C_4
Trinidad & Tobago	C_2	0.731	C_2	Mali	C_4	0.377	C_4
Malaysia	C_2	0.729	C_2	Mauritania	C_4	0.377	C_4
Mauritius	C_2	0.698	C_2	Nigeria	C_4	0.374	C_4
Brazil	C_2	0.672	C_2	Togo	C_4	0.355	C_4
u_2		0.672		Cameroon	C_4	0.354	C_4
Thailand	C_3	0.639	C_3	Malawi	C_4	0.350	C_4
Costa Rica	C_3	0.623	C_3	Nicaragua	C_4	0.206	C_4
Peru	C_3	0.584	C_3				

Results of the UTADIS I Method

The different objective between UTADIS I and UTADIS leads to results that differ from the corresponding results obtained through the UTADIS method, although the classification accuracy is once again 100.

The global utilities of the high-income economies are very close to 1 (most of the global utilities are over 0.999), so that the distance from the utility threshold (0.4784) is maximized. Only Portugal's global utility is close to the utility threshold. The rest of the high-income economies obtain global utilities over 0.7102. Concerning the upper-middle income economies, Greece and South Korea were found

Table 3.2 Classification results obtained through the UTADIS I method

Countries	Estimated class	Utility	Actual class	Countries	Estimated class	Utility	Actual class
Switzerland	C_1	1.000	C_1	Thailand	C_3	0.204	C_3
Norway	C_1	1.000	C_1	Peru	C_3	0.178	C_3
Japan	C_1	1.000	C_1	Tunisia	C_3	0.151	C_3
Netherlands	C_1	1.000	C_1	Algeria	C_3	0.138	C_3
Belgium	C_1	1.000	C_1	Jordan	C_3	0.120	C_3
France	C_1	1.000	C_1	El Salvador	C_3	0.113	C_3
United Kingdom	C_1	1.000	C_1	Dominican Republic	C_3	0.111	C_3
Sweden	C_1	1.000	C_1	Romania	C_3	0.106	C_3
Austria	C_1	1.000	C_1	Ecuador	C_3	0.106	C_3
Finland	C_1	1.000	C_1	Guatemala	C_3	0.099	C_3
United States	C_1	1.000	C_1	Morocco	C_3	0.095	C_3
Denmark	C_1	1.000	C_1	Papua New Guinea	C_3	0.094	C_3
Italy	C_1	1.000	C_1	Philippines	C_3	0.078	C_3
Australia	C_1	1.000	C_1	Indonesia	C_3	0.068	C_3
Canada	C_1	1.000	C_1	Bolivia	C_3	0.052	C_3
Israel	C_1	0.823	C_1	u_3		0.052	
Ireland	C_1	0.728	C_1	Egypt	C_4	0.051	C_4
Spain	C_1	0.719	C_1	Sri Lanka	C_4	0.051	C_4
New Zealand	C_1	0.710	C_1	Cameroon	C_4	0.045	C_4
Portugal	C_1	0.479	C_1	China	C_4	0.043	C_4
u_1		0.478		Ivory Coast	C_4	0.038	C_4
Greece	C_2	0.477	C_2	Senegal	C_4	0.038	C_4
Korea, Rep.	C_2	0.477	C_2	Mauritania	C_4	0.027	C_4
Uruguay	C_2	0.398	C_2	Nicaragua	C_4	0.024	C_4
Mexico	C_2	0.356	C_2	Pakistan	C_4	0.024	C_4
Hungary	C_2	0.327	C_2	Ghana	C_4	0.021	C_4
Trinidad & Tobago	C_2	0.319	C_2	India	C_4	0.016	C_4
Chile	C_2	0.300	C_2	Togo	C_4	0.013	C_4
Malaysia	C_2	0.296	C_2	Nigeria	C_4	0.010	C_4
Mauritius	C_2	0.267	C_2	Kenya	C_4	0.008	C_4
Brazil	C_2	0.253	C_2	Mali	C_4	0.007	C_4
u_2		0.253		Bangladesh	C_4	0.006	C_4
Turkey	C_3	0.212	C_3	Nepal	C_4	0.004	C_4
Poland	C_3	0.204	C_3	Malawi	C_4	0.000	C_4
Costa Rica	C_3	0.204	C_3				

to be the most creditworthy and economically sound countries in this group. This result was also obtained through the UTADIS method. Turkey, Poland, Costa Rica, and Thailand were found to be the less risky countries within the group of lower-middle income economies, whereas Malawi, Nepal, and Bangladesh are the most risky countries. Table 3.2 presents in detail the obtained results, as well as the original and the estimated classification of the countries.

Concerning the significance of the evaluation criteria in the sorting model developed through the UTADIS I method, the GNP per capita is clearly the dominant factor with a weight of 98.62%. This extreme weight is fully justifiable, as it has been already observed, in this specific case study, that the GNP per capita is able by itself to provide an accurate classification of the countries in the four classes of risk defined by the World Bank. The weights of the other evaluation criteria are less than 1%.

Results of the UTADIS II Method

The additive utility model developed by the UTADIS II method is similar to the model that was developed through the UTADIS method. This new model is also able to provide a correct classification of the countries to the classes to which they belong, providing a classification accuracy of 100%. The global utilities of the countries are a little bit lower than the global utilities that were calculated through the UTADIS method, but they are still similar. Switzerland was once again found to be the most creditworthy country with global utility of 0.9692, followed by Japan (with global utility 0.9579), United States (with global utility 0.9566), The Netherlands (with global utility 0.9529), and the United Kingdom (with global utility 0.9504). On the contrary, Nicaragua was found the be the most risky country (with global utility 0.2026), followed by Malawi (with global utility 0.2866) and Togo (with global utility 0.3309).

The significance of the evaluation criteria is also very similar to corresponding results obtained through the UTADIS method. More specifically, the GNP per capita is once again the most important criterion for the evaluation of country risk. Its weight is the same with the weight that was estimated using the UTADIS method (52.14%). Among the rest of the evaluation criteria, the most significant ones were found to be the current account balance as percentage of GNP and the imports' average annual growth rate with weights 8.71% and 7.39%, respectively.

Results of the UTADIS III Method

The final method that is applied in this case study in order to assess country risk is the UTADIS III method. This method's objective is to minimize the number of misclassification and at the same time to maximize the distances of the correctly classified countries from the utility thresholds.

As in all of the previous variants of the UTADIS method, the developed additive utility model developed through the UTADIS III method is also able to correctly classify the 66 countries under consideration to their original class. The obtained results, as expected, are very similar to the results of the UTADIS I method. There are some small differences in the global utilities of the countries, but overall, both the global utilities as well as the marginal utilities of the evaluation criteria are very similar. In this new model, the weight of the GNP per capita is 98.78%. Only two other criteria are considered by the model; the life expectancy and the current account balance as percentage of GNP, with weights of 1.11% and 0.11%, respectively. All the other evaluation criteria are not included in this model.

Table 3.3 Standardized canonical discriminant function coefficients

	Function 1	Function 2	Function 3
Current account balance as percentage of GNP	0.4589	–0.0446	–0.1625
Exports average annual growth rate	0.1163	0.3130	0.1023
Imports average annual growth rate	0.2232	–0.2516	0.0878
GNP per capita	–0.5241	0.5791	–0.2487
Average annual growth rate of GNP per capita	0.1601	0.2597	0.7263
Gross domestic investment	0.0259	–0.0178	0.5867
External debt as percentage of GNP	–0.4666	–0.6644	–0.2980
Gross international reserves as percentage of GNP	0.5690	0.3161	0.9879
Reserves to imports ratio	–0.0029	–0.2309	–1.2547
Net foreign debt to exports ratio	–0.5325	–0.0567	0.5534
Life expectancy	0.3090	–0.0154	0.3111
Political risk	0.1828	0.1854	–0.6054

Comparison with Discriminant Analysis

For comparison purposes, discriminant analysis was also applied in the sample of countries under consideration in order to develop a discriminant model to classify the countries in their original class. Discriminant analysis is a well-known multivariate statistical method for the study of classification problems. The objective of performing the discriminant analysis was to examine how a different statistical approach could perform in this specific case study compared with the UTADIS method and its variants. Using the discriminant analysis, three discriminant functions were developed. The standardized canonical discriminant function coefficients are presented in Table 3.3.

The developed discriminant model based on these three discriminant functions is unable to correctly classify all countries in their original class. More specifically, there are 9 misclassified countries, resulting in an overall classification accuracy of 86.36%. On the contrary, as already presented, the UTADIS, I, II, and III methods were all able to provide an accurate assignment of each country to its original (predefined) class (classification accuracy 100%). This result clearly depicts the superiority of the preference disaggregation approach over the discriminant analysis, at least in this specific case study.

A detailed error analysis of the results obtained by the discriminant analysis is presented in Table 3.4. The first part of Table 3.4 presents how the classification of the countries was made by the discriminant functions. The diagonal represents the correct classifications, and all the other elements represent the differences (misclassifications) between the actual classification of the countries and their classification by the discriminant functions. The second part of Table 3.4 presents the same information expressed as percentage of the number of countries that are included in each original class.

Table 3.4 Error summary of the classification results obtained by discriminant analysis

		Estimated class							
		C_1	C_2	C_3	C_4	C_1	C_2	C_3	C_4
	C_1	19	1	–	–	95.0%	5.0%	–	–
Original	C_2	–	8	2	–	–	80.0%	20.0%	–
class	C_3	–	2	15	1	–	11.10%	83.3%	5.6%
	C_4	–	–	3	15	–	–	16.7%	83.3%

Results of the UTASTAR Method

The additive utility model that was developed using the ranking provided by Euromoney was unable to represent consistently the evaluation policy of the officers of Euromoney. More specifically, the ranking obtained by the developed additive utility model depicted some differences with the initial ranking of Euromoney. These inconsistencies are justified by the differences between the evaluations provided by Euromoney and the corresponding estimations of World Bank. Nevertheless, the inconsistencies of the developed country risk model are not considered to be significant ones. The most significant inconsistency concerns Nepal. According to Euromoney's country risk rating, Nepal was ranked in the 57th place among the 66 countries of this case study. The additive utility model that was developed through the UTASTAR method ranks Nepal in the 50th place. However, it should be noted that the difference between the global utility of Nepal (0.316) and the global utility of Senegal (0.308), which is ranked in the 57th place, is small. Consequently, although Nepal is ranked higher by the developed model, its score (global utility) is still similar to other under-developed countries.

The similarity between the rankings of Euromoney and the model is also confirmed using the Kendall's τ rank correlation coefficient. The value of Kendall's τ is 0.961, very close to 1, showing that there is a significant consistency between the two rankings. Figure 3.1 illustrates the countries' ranking versus their global utilities estimated by the UTASTAR method.

According to the global utilities of the 66 countries under consideration, three major groups can be distinguished. The first one includes 22 countries whose global utility is over 0.8. This group includes all the high-income economies, except Israel, and three countries that are considered by World Bank as upper-middle economies (South Korea, Malaysia, and Chile). The second group includes 18 countries with global utilities ranging between 0.425 and 0.479. Most of these countries are in the upper-middle income group. Additionally, this group also includes Israel (high-income economy), Thailand, Indonesia, Philippines, Poland, Morocco, and Turkey, which are considered by World Bank as lower-middle income economies. Finally, the third group consists of 26 countries with global utilities below 0.35. These countries are considered as lower-middle and low-income economies. In this group of risky countries, the most untrustworthy one was found to be Nicaragua, with global utility 0.133.

The significance of the evaluation criteria in the ranking model developed through the UTASTAR method differs from the importance of the evaluation criteria in

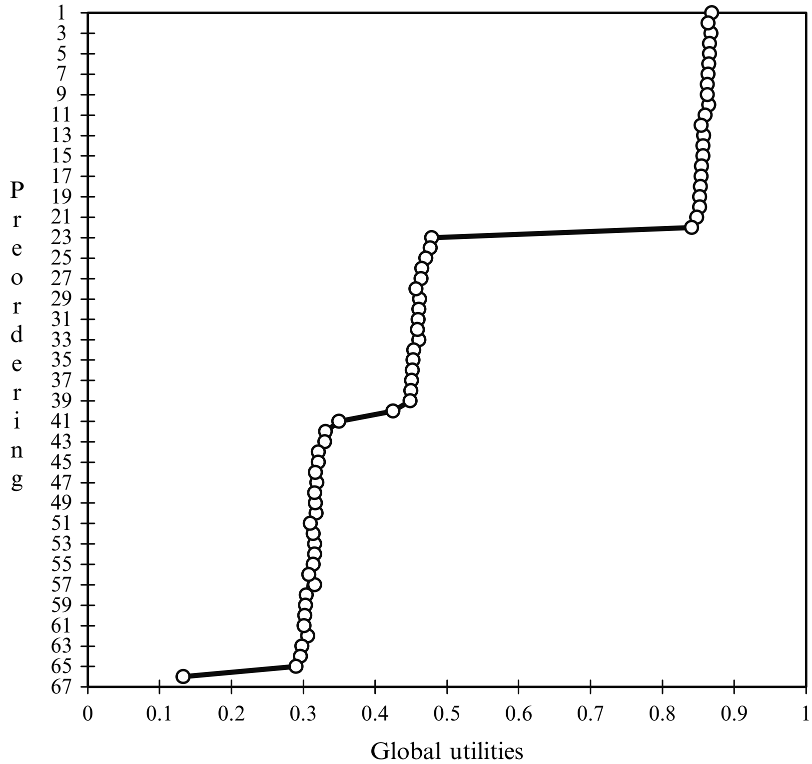

Fig. 3.1 Countries' ranking versus global utilities

the classification models developed through the UTADIS method and its variants. More specifically, the most important criterion in the ranking model developed by the UTASTAR method is the political risk followed by the gross domestic investment, the net foreign debt/exports, and the import's average annual growth rate, with weights 55.45%, 12.11%, 11.96%, and 11.93%, respectively. This result is in accordance with the decision policy of Euromoney's managers, who considers political risk as the most important criterion in their country risk rating.

3.2 The Study of Doumpos and Zopounidis (2000)

The performance of the MHDIS method and its applicability in financial risk assessment are explored in this case through an application to country risk evaluation. The recent economic crises have demonstrated in the clearest way that country risk is a crucial risk factor with significant impact on any corporate entity with an international activity. This significance of the country risk assessment problem, along with its complexity that is due to the plethora of factors of different nature that are

Table 3.5 Economic indicators (evaluation criteria)

g_1	Current account balance/GDP
g_2	Export volume growth
g_3	Gross domestic investment/GDP
g_4	Import volume growth
g_5	Inflation (GDP deflator)
g_6	Net trade in goods and services
g_7	Present value of debt/Exports of goods and services
g_8	Present value of debt/GNP
g_9	Total debt service/GNP
g_{10}	Income velocity of money (GDP/M2)
g_{11}	GNP growth
g_{12}	Gross international reserves in months of imports

involved (e.g., macroeconomic, social, political factors, etc.), make country risk assessment a challenging research problem where several scientific fields such as statistical analysis and operations research can provide significant contribution. This is the main reason that justifies the selection of country risk assessment as a field that is appropriate to examine the applicability of the MHDIS method in financial risk assessment.

3.2.1 Data Set Description

The sample used in this application is derived from the World Bank (World Bank, 1997). The data refer to 143 countries for the year 1995. They involve a significantly large number of indicators and variables relative to country risk assessment including inflation and exchange rates, the balance of payments, tax policies, macroeconomic indicators, indicators upon structural transformation, as well as trade indicators, external debt indicators, etc. (98 indicators overall; Pentaraki et al., 1999). Obviously, the incorporation in the analysis of such a large number of evaluation criteria would result in the development of an unrealistic country risk assessment model with limited practical value. To overcome this problem, a factor analysis is performed to select the most relevant criteria that best describe the economic performance and the creditworthiness of the countries. It could be possible to override factor analysis if a country risk expert was available to determine the most significant country risk indicators, or if the decision maker had a clear view of the indicators that should be examined. Nevertheless, in any case the factor analysis results provide significant support in determining the indicators that characterize the economic performance and the creditworthiness of a country. On the basis of the factor analysis results (i.e., factor loadings) and the relevance of the considered criteria to country risk assessment as reported in the international literature (Saini and Bates, 1984; Cosset et al. 1992; Oral et al., 1992), 12 evaluation criteria are finally selected to be included in the developed country risk assessment model (Table 3.5).

According to the World Bank, the countries under consideration are categorized into four classes according to their income level: (1) High-income economies (class C_1) including 31 countries, mostly European ones, as well as the United States, Canada, Australia, New Zealand, Japan, Hong Kong, etc. (2) Upper-middle economies (class C_2), including 21 countries from Europe (e.g., Greece and Hungary), South and Eastern Asian, and Latin and South America. (3) Lower-middle income economies (class C_3), including 42 countries from Eastern Europe, Asia, Africa, and South and Latin America. (4) Low-income economies (class C_4), including 49 countries, mostly from Africa and Asia. This classification constitutes the basis for the development of the appropriate country risk assessment model using the MHDIS method.

3.2.2 Illustration of MHDIS on the Complete Sample

Because the sample used involves four classes of countries, the hierarchical discrimination process of the MHDIS method that was described in Chapter 2 consists of three stages. In the first stage, the discrimination among the countries belonging in the high-income economy group and the countries belonging in the rest of the classes is performed. In the second stage, the countries belonging to the upper-middle income economy group are discriminated from the countries of the lower-middle and the low-income economy groups. Finally, the third stage involves the discrimination among the countries of the lower-middle and the low-income economy groups.

Each of these three stages involves the solution of the three mathematical programming problems that have already been described (LP1, MIP, and LP2). In the first stage, there are no misclassifications. The pair of additive utility functions that are developed is able to discriminate accurately the countries belonging in the first group (high-income economy) from the rest of the countries. In the second stage, solving LP1 results in one misclassified country, China, which is classified into the group of upper-middle income economies while belonging in the group of low-income economies. Because there is only one misclassification, MIP is not solved, and the procedure proceeds with the solution of LP2 (China is retained as misclassified). China is ignored during the third stage of the hierarchical discrimination process, after which no additional classification errors are encountered. The additive utility functions developed at each stage of the hierarchical discrimination process along with the resulting classification rule are presented below.

Stage 1

$$
\begin{aligned}
U_1(\mathbf{g}) = {} & 0.0093u_{1,1}(g_1) + 0.0137u_{1,2}(g_2) + 0.0095u_{1,3}(g_3) \\
& + 0.02u_{1,4}(g_4) + 0.0127u_{1,5}(g_5) + 0.0112u_{1,6}(g_6) \\
& + 0.8668u_{1,7}(g_7) + 0.0099u_{1,8}(g_8) + 0.0103u_{1,9}(g_9) \\
& + 0.0113u_{1,10}(g_{10}) + 0.0136u_{1,11}(g_{11}) + 0.0117u_{1,12}(g_{12})
\end{aligned}
$$

$$\begin{aligned}U_{\sim 1}(\mathbf{g}) = {} & 0.0894u_{\sim 1,1}(g_1)+0.0137u_{\sim 1,2}(g_2)+0.0095u_{\sim 1,3}(g_3)\\ &+0.0113u_{\sim 1,4}(g_4)+0.0592u_{\sim 1,5}(g_5)+0.0112u_{\sim 1,6}(g_6)\\ &+0.7489u_{\sim 1,7}(g_7)+0.0099u_{\sim 1,8}(g_8)+0.0103u_{\sim 1,9}(g_9)\\ &+0.0113u_{\sim 1,10}(g_{10})+0.0136u_{\sim 1,11}(g_{11})+0.0117u_{\sim 1,12}(g_{12})\end{aligned}$$

Stage 2

$$\begin{aligned}U_2(\mathbf{g}) = {} & 0.0068u_{2,1}(g_1)+0.0602u_{2,2}(g_2)+0.2635u_{2,3}(g_3)\\ &+0.0087u_{2,4}(g_4)+0.0576u_{2,5}(g_5)+0.2099u_{2,6}(g_6)\\ &+0.0167u_{2,7}(g_7)+0.0098u_{2,8}(g_8)+0.0102u_{2,9}(g_9)\\ &+0.1263u_{2,10}(g_{10})+0.0107u_{2,11}(g_{11})+0.2195u_{2,12}(g_{12})\\ U_{\sim 2}(\mathbf{g}) = {} & 0.1086u_{\sim 2,1}(g_1)+0.0251u_{\sim 2,2}(g_2)+0.0069u_{\sim 2,3}(g_3)\\ &+0.1116u_{\sim 2,4}(g_4)+0.0261u_{\sim 2,5}(g_5)+0.2502u_{\sim 2,6}(g_6)\\ &+0.0157u_{\sim 2,7}(g_7)+0.0641u_{\sim 2,8}(g_8)+0.0102u_{\sim 2,9}(g_9)\\ &+0.2128u_{\sim 2,10}(g_{10})+0.1171u_{\sim 2,11}(g_{11})+0.0516u_{\sim 2,12}(g_{12})\end{aligned}$$

Stage 3

$$\begin{aligned}U_3(\mathbf{g}) = {} & 0.0052u_{3,1}(g_1)+0.1806u_{3,2}(g_2)+0.3212u_{3,3}(g_3)\\ &+0.0068u_{3,4}(g_4)+0.008u_{3,5}(g_5)+0.1627u_{3,6}(g_6)\\ &+0.1183u_{3,7}(g_7)+0.0082u_{3,8}(g_8)+0.0085u_{3,9}(g_9)\\ &+0.0757u_{3,10}(g_{10})+0.0086u_{3,11}(g_{11})+0.0963u_{3,12}(g_{12})\\ U_{\sim 3}(\mathbf{g}) = {} & 0.2597u_{\sim 3,1}(g_1)+0.0085u_{\sim 3,2}(g_2)+0.1386u_{\sim 3,3}(g_3)\\ &+0.0068u_{\sim 3,4}(g_4)+0.008u_{\sim 3,5}(g_5)+0.1949u_{\sim 3,6}(g_6)\\ &+0.0943u_{\sim 3,7}(g_7)+0.0082u_{\sim 3,8}(g_8)+0.0085u_{\sim 3,9}(g_9)\\ &+0.1222u_{\sim 3,10}(g_{10})+0.1187u_{\sim 3,11}(g_{11})+0.0313u_{\sim 3,12}(g_{12})\end{aligned}$$

Classification rule

If $U_1(\mathbf{g}_1) > U_{\sim 1}(\mathbf{g}_1)$ then country $\mathbf{x}_j$ is classified in the high-income group
Else if $U_2(\mathbf{g}_1) > U_{\sim 2}(\mathbf{g}_1)$ then country $\mathbf{x}_j$ is classified in the upper-middle income group
Else if $U_3(\mathbf{g}_1) > U_{\sim 3}(\mathbf{g}_1)$ then country $\mathbf{x}_j$ is classified in the lower-middle income group
Else country $\mathbf{x}_j$ is classified in the low-income group

The overall classification accuracy of this country risk evaluation model is 99.30% (1 misclassified country: China). An example of the form of the marginal utility functions in this classification model is presented in Figure 3.2 for criterion gross domestic investment/GDP (g_3), which is among the most significant criteria in all additive utility functions that are developed (except for the discrimination

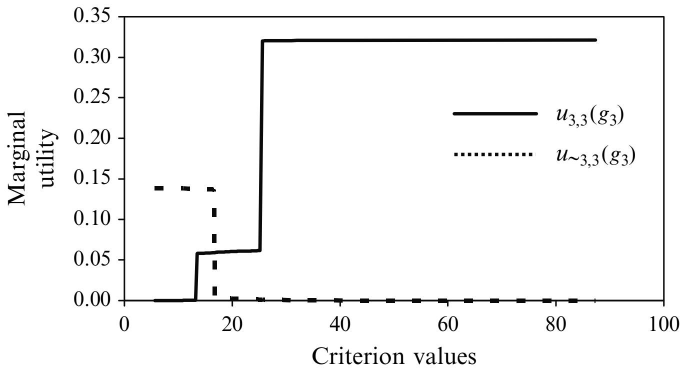

Fig. 3.2 Marginal utilities of criterion “gross domestic investment/GDP” in discriminating between upper-middle income and low-income economies

among high-income economies from the rest of the countries). Figure 3.2 illustrates the marginal utilities of this criterion in the additive utility functions U_3 and $U_{\sim 3}$ that discriminate between the lower-middle income group and the low-income group.

3.2.3 Validation Tests

Whereas the above discussion is just illustrative of the functionality of the MHDIS method in discriminating a set of alternatives (countries) into a predefined set of classes, it does not provide any insight on the generalizing performance of the developed models in classifying alternatives not included in the training sample.

To perform this analysis, a series of validation tests is conducted as follows. The complete sample of 143 countries is split, at random, into two subsamples. The first one consisting of 102 countries is used as the training sample, and the second sample consisting of 41 countries is used to examine the performance of the developed model in classifying countries that are not considered in the training sample. Henceforth, this second sample will be referred to as the validation sample. The proportions of the four groups of countries in the validation sample are set to be equivalent to the proportions of the groups in the complete sample. For instance, in the complete sample the high-income group includes 31 countries, which account for 21.68% of the total sample. To respect this proportion, the validation sample includes 9 countries of the first class, which account for 21.95% of the total number of countries in the validation sample. The number of countries belonging into the other three groups in the validation sample is determined in a similar way. Thus, the upper-middle income group includes 6 countries, the lower-middle income group includes 12 countries, and the low-income group includes 14 countries. All countries included in the validation sample are randomly selected, with the only restriction being the proportions of the groups as described above. Overall, 40 validation tests are performed in order to have as much of an unbiased estimate of the performance of the MHDIS method as possible.

Table 3.6 Average classification results of MHDIS and MDA over the 40 validation tests

		Training sample				
		Estimated classification				Overall
		C_1	C_2	C_3	C_4	accuracy
MHDIS	C_1	100.00%	0.00%	0.00%	0.00%	99.07%
	C_2	0.00%	98.83%	1.17%	0.00%	
	C_3	0.00%	0.00%	100.00%	0.00%	
	C_4	0.00%	1.50%	0.71%	97.79%	
MDA	C_1	91.48%	1.14%	8.07%	0.34%	79.31%
	C_2	6.50%	47.50%	41.17%	4.83%	
	C_3	6.67%	6.75%	81.17%	5.42%	
	C_4	0.07%	2.00%	14.21%	83.71%	
		Validation sample				
MHDIS	C_1	100.00%	0.00%	0.00%	0.00%	75.55%
	C_2	0.00%	44.17%	38.75%	17.08%	
	C_3	0.00%	23.75%	62.50%	13.75%	
	C_4	0.00%	5.36%	10.18%	84.46%	
MDA	C_1	89.17%	0.83%	10.00%	0.00%	70.73%
	C_2	8.75%	26.25%	52.08%	12.92%	
	C_3	9.79%	9.79%	75.00%	6.46%	
	C_4	0.36%	7.86%	17.50%	74.29%	

For comparison purposes, discriminant analysis (DA) is also applied in these validation tests. DA can be considered as the first approach to introduce multiple factors (variables) in the discrimination among different groups of objects. When there are more than two groups, the application of multiple discriminant analysis (MDA) leads to the development of linear discriminant functions that maximize the ratio of among group to within group variability; this assumes that the variables follow a multivariate normal distribution and that the dispersion matrices of the groups are equal. In this case study, MDA is selected for comparison purposes due to its popularity in the field of finance in studying financial decision problems requiring a grouping of a set of alternatives (Altman et al., 1981). Furthermore, the method is popular among academic researchers in evaluating the performance of new algorithms and approaches to study general classification problems. Finally, it should be also noted that MDA has already been applied in several studies on country risk assessment (Saini and Bates, 1984).

The classification results of MHDIS and MDA on the 40 validation tests are presented in Table 3.6. All figures reported in this table correspond with averages over all validation tests. The elements $C_1 - C_1$, $C_2 - C_2$, $C_3 - C_3$, and $C_4 - C_4$ represent the average classification accuracy for each of the four classes, and all other elements correspond with average classification errors.

With regard to the training sample, the overall classification accuracy of MHDIS is significantly higher than the one of MDA (99.07% for MHDIS vs. 79.31% for MDA). This difference is not surprising bearing in mind two facts: (1) MHDIS results in the development of piecewise linear discrimination models as opposed to the linear discrimination of MDA, (2) the discrimination models developed through MHDIS involve more degrees of freedom compared with MDA, thus ensuring better model fit.

Of course, higher model fit does not ensure higher generalizing ability, which is the ultimate objective in decision models developed through regression-based techniques. In that respect, the results on the validation tests are of particular interest toward the evaluation of the predictability of both MHDIS and MDA. The results presented in Table 3.6 indicate that in terms of the overall classification accuracy, MHDIS still performs better than MDA, although the difference among the two methods is smaller compared with the training sample. The average overall classification accuracy of the MHDIS method over the 40 replications is 75.55%, while the overall classification accuracy of MDA is 70.73%. The t-value of a t-test regarding the difference in these average classification accuracies is 4.37. Thus, the difference between the two methods is significant at the 1% level.

In terms of the individual classification accuracies in the validation tests, MHDIS always classifies correctly all countries belonging in the high-income group (class C_1). Furthermore, there is no country belonging in the other classes that is classified incorrectly into the high income group. MDA also performs quite well in identifying high-income economies. However, on average 10% of the countries actually belonging into the high-income group are classified as lower-middle income economies (class C_3), whereas a small percentage (0.36%) of countries belonging into the low-income group (class C_4) are assigned by MDA into the high-income group; this is a significant misclassification. Similar to the high income group, both MHDIS and MDA perform quite well in identifying countries of the low-income group. The average classification accuracy for this group is 84.46% for MHDIS and 74.29% for MDA. MDA also performs satisfactory in the lower-middle income group (classification accuracy 75%) as opposed to the 62.5% accuracy of the MHDIS method. Finally, it is apparent that the major problem in both methods is to identify the countries belonging in the upper-middle income group. MHDIS's average classification accuracy for this group is 44.17%, as opposed to 26.25% for MDA. However, it should be pointed out that most of the upper-middle income economy countries that are misclassified by both methods are assigned to the group of lower-middle income economies.

3.3 The Study of Doumpos and Zopounidis (2002)

Similary to the previous application, in this application the MHDIS method is applied to country risk data obtained from the World Bank (World Bank, 2002). As it was already mentioned, the data provided in the World Bank's tables span the period 1960–2000. For the present analysis, the period 1996–2000 was selected. On the basis of data availability for this period, finally 161 countries were selected out of the total 207 countries included in the tables. On the basis of the most recent data (year 2000), the World Bank classifies the countries in four income-level groups: (1) high-income economies (group C_1), (2) upper-middle income economies (group C_2), (3) lower-middle income economies (group C_3), and (4) low-income

Table 3.7 Country risk indicators

g_1	Current account balance/GDP
g_2	Exports of goods and services/GDP
g_3	Foreign direct investment/GDP
g_4	Annual GDP growth
g_5	Annual GDP per capita growth
g_6	Gross capital formation/GDP
g_7	Gross fixed capital formation/GDP
g_8	Imports of goods and services/GDP
g_9	Inflation
g_{10}	Infant mortality rate
g_{11}	Short-term debt/Total external debt
g_{12}	Total debt service/Exports of goods and services
g_{13}	Total debt service/GNI
g_{14}	Life expectancy at birth
g_{15}	Consumer price index
g_{16}	Annual growth of exports of goods and services
g_{17}	Net domestic credit/GDP
g_{18}	Total external debt /Exports of goods and services
g_{19}	Total external debt/GDP
g_{20}	Total debt service/Imports of goods and services
g_{21}	Total debt service/Gross international reserves
g_{22}	Gross international reserves in months of imports

economies (group C_4). This classification constitutes the basis for the analysis in this study, assuming the income level group into which a country is assigned to is closely related to its overall performance and risk. Of course, the overall performance and risk of a country cannot be fully represented in this four-group classification, but nevertheless, the income level classification of the World Bank can be considered as an adequate proxy for the development, performance, and risk of the countries.

An initial preprocessing of the available data led to the selection of 22 possible country risk indicators (Table 3.7), selected according to (1) their relevance to country risk analysis on the basis of the existing literature on the subject (Saini and Bates, 1984; Zopounidis et al., 1998), and (2) the availability of the data (for many other relevant indicators, there were too many missing data in the World Bank's tables). The selected indicators involve, mainly, the economic performance of the countries (debt, trade, inflation, GDP, etc.), but there are also indicators involving some social aspects of the problem such as the life expectancy at birth and the infant mortality rate, which are closely related to the quality of living and the development of each country. Table 3.8 presents a statistical analysis (ANOVA) for the selected ratios, with regard to their statistical significance in discriminating the four aforementioned groups of countries in the sample.

The ANOVA results of Table 3.8 indicate that there are 15 indicators that statistically differentiate the four groups in at least one of the years (at the 1% significance level), whereas 10 indicators statistically differentiate the four groups in all five years of the analysis.

Table 3.8 ANOVA results for the selected country risk indicators (F statistic)

	1996	1997	1998	1999	2000
g_1	3.69	10.98*	8.86*	5.94*	4.47*
g_2	6.80*	5.93*	6.71*	5.53*	9.97*
g_3	0.76	0.86	1.01	1.73	6.12*
g_4	0.92	0.14	0.03	1.44	0.50
g_5	1.00	0.72	0.77	1.59	2.66
g_6	1.72	4.23*	3.41	1.36	2.01
g_7	1.44	3.28	2.34	1.03	1.41
g_8	0.71	0.58	0.82	0.43	2.86
g_9	0.25	0.44	4.33*	1.56	0.63
g_{10}	72.34*	72.12*	69.66*	69.21*	68.12*
g_{11}	20.08*	28.46*	31.05*	27.85*	25.15*
g_{12}	16.61*	18.29*	18.00*	13.49*	15.14*
g_{13}	12.29*	16.34*	13.73*	16.03*	21.57*
g_{14}	89.92*	83.74*	77.21*	68.93*	66.05*
g_{15}	0.25	0.14	0.60	0.59	0.53
g_{16}	192.33*	211.02*	234.56*	250.92*	225.79*
g_{17}	1.16	0.39	1.17	0.20	0.47
g_{18}	38.83*	38.33*	21.75*	31.36*	27.22*
g_{19}	17.32*	20.73*	17.80*	26.91*	26.06*
g_{20}	23.81*	23.96*	23.19*	23.51*	22.87*
g_{21}	14.90*	16.54*	17.38*	13.84*	17.57*
g_{22}	1.18	0.91	0.79	0.75	12.45*

On the basis of these results, it was decided to proceed with the development of country risk assessment models at two levels. In the first case, all the 22 country risk indicators are considered, whereas in the second case a reduced set of 12 indicators is employed. These 12 indicators are selected among the 15 indicators that were found to statistically differentiate the four groups of countries for at least one year of the analysis, excluding some indicators that had high correlations with others (correlation coefficient higher than 0.7, in absolute terms). In particular, high correlations were observed for the pairs of indicators: (1) infant mortality rate (g_{10}) and life expectancy at birth (g_{14}) with correlation coefficient -0.94, (2) total debt service/exports of goods and services (g_{12}) and total debt service/imports of goods and services (g_{20}) with correlation coefficient 0.90, and (3) total external debt/exports of goods and services (g_{18}) and total external debt/GDP (g_{19}) with correlation coefficient 0.76. From these pairs of correlated indicators, it was decided to retain in the analysis the indicators that are most relevant to country risk assessment, on the basis of the existing literature (Zopounidis et al., 1998), namely infant mortality rate (g_{10}), total debt service/exports of goods and services (g_{12}), and total external debt/GDP (g_{19}).

Of course, within the context of MCDA, such a statistical analysis for the selection of the decision criteria (country risk indicators) is not considered to be necessary, given that a decision maker can specify the criteria that he considers relevant to the analysis. However, in this case study such an expert decision maker was not available. On the other hand, the consideration of a large number of criteria (the complete set of 22 indicators) may pose practical problems in the implementation of the development model, mainly with regard to the requirement that a significant volume of data needs to be collected to implement the model. Because data collection is often a time consuming and possibly costly process, it was decided in this analysis to explore the possibility of using a reduced set of indicators selected through the aforementioned approach.

3.3.1 Presentation of Results

The development and testing of country risk assessment models through the MHDIS method is performed through the following methodology. Initially, the data of the most recent year (year 2000) are used as the reference set for the development of an appropriate country risk model. This model is then applied in all the previous four years of the analysis (years 1996–1999) to evaluate its ability in providing early warning signals for the current situation of the countries.

The country risk assessment model developed through this approach consists of six additive utility functions. The utility functions U_1 and $U_{\sim 1}$ are used to decide whether a country belongs in the high income group (C_1) or not. For the countries not assigned into group C_1, the second pair of utility functions U_2 and $U_{\sim 2}$ is employed to decide whether they should be classified into the upper-middle income group (C_2) or not. Finally, the third pair of utility functions U_3 and $U_{\sim 3}$ is employed for the countries not assigned in the previous stage into group C_2, to decide if they should be classified as lower middle income economies (group C_3) or as low income economies (group C_4). Therefore, the classification of a country $\mathbf{x}_j$ into one for the four groups is performed through the following rules:

If $U_1(\mathbf{g}_j) > U_{\sim 1}(\mathbf{g}_j)$ then $\mathbf{x}_j \in C_1$
else if $U_2(\mathbf{g}_j) > U_{\sim 2}(\mathbf{g}_j)$ then $\mathbf{x}_j \in C_2$
else if $U_3(\mathbf{g}_j) > U_{\sim 3}(\mathbf{g}_j)$ then $\mathbf{x}_j \in C_3$
else $\mathbf{x}_j \in C_4$

The contributions (weights) of the considered country risk indicators in the developed additive utility functions are illustrated in Table 3.9. The presented results involve two country risk models. Model A is developed considering the complete set of country risk indicators of Table 3.7, whereas model B is based on a limited set of 12 indicators. In model A, there are nine indicators with weight higher than 10% in at least one of the six utility functions (indicators g_1, g_2, g_3, g_8, g_{14}, g_{17}, g_{19}, g_{21}, g_{22}), whereas in model B all the indicators have weight higher than 10% in at least one of the six utility functions, except for the indicator g_{11} (short-term debt/total external debt). For the indicators used in both models, the main differences between models A and B involve the significance given to indicators g_6, g_9, g_{10}, g_{11}, g_{12}, and g_{13}. These six indicators do not contribute significantly (weight lower than 10%) in any of the utility functions of model A, whereas they are found significant in at least one of the utility functions of model B. On the other hand, both models agree

Table 3.9 Weights of the country risk indicators in the MHDIS models (in %)

	Model A (complete set of indicators)						Model B (reduced set of 12 indicators)					
	U_1	$U_{\sim 1}$	U_2	$U_{\sim 2}$	U_3	$U_{\sim 3}$	U_1	$U_{\sim 1}$	U_2	$U_{\sim 2}$	U_3	$U_{\sim 3}$
g_1	7.03	28.25	0.10	10.78	0.79	1.85	16.55	0.78	6.57	28.06	22.36	2.19
g_2	0.03	7.17	9.00	9.75	14.05	12.91	0.80	0.80	0.65	0.65	11.28	8.11
g_3	0.03	0.03	27.84	7.02	3.81	7.02	10.88	0.73	0.59	9.69	0.76	3.11
g_4	0.03	0.03	0.02	0.02	0.02	0.70	–	–	–	–	–	–
g_5	0.03	3.52	0.02	0.02	0.02	0.02	–	–	–	–	–	–
g_6	3.44	3.28	3.36	1.38	3.94	0.02	17.44	0.79	12.58	0.64	3.49	0.50
g_7	0.44	0.03	4.27	4.38	1.45	0.02	–	–	–	–	–	–
g_8	7.03	0.03	10.80	24.48	11.88	5.02	–	–	–	–	–	–
g_9	0.03	0.20	0.03	0.03	7.02	1.65	4.21	13.06	2.54	5.66	5.74	0.51
g_{10}	0.03	3.25	7.04	2.66	7.30	8.91	0.77	47.95	26.25	19.84	11.33	20.49
g_{11}	0.02	3.36	0.02	0.02	0.74	1.06	8.75	0.61	0.61	0.61	0.49	0.49
g_{12}	0.02	3.28	0.02	1.30	7.02	1.80	15.38	2.67	7.76	1.79	0.49	11.71
g_{13}	0.02	7.02	0.02	0.02	3.09	1.89	0.60	0.60	6.93	0.60	12.03	0.49
g_{14}	41.00	15.04	7.85	6.46	0.95	7.02	–	–	–	–	–	–
g_{15}	0.03	0.03	0.03	2.58	0.02	0.02	–	–	–	–	–	–
g_{16}	6.64	0.03	0.03	0.03	0.02	0.02	–	–	–	–	–	–
g_{17}	10.04	4.00	1.46	15.75	15.09	14.65	6.60	24.20	33.90	4.29	6.99	29.10
g_{18}	7.02	6.84	5.09	0.02	7.02	7.02	–	–	–	–	–	–
g_{19}	0.02	0.02	12.27	1.66	8.14	5.51	0.42	6.73	1.20	27.80	16.12	16.21
g_{20}	0.01	7.01	0.01	0.01	7.01	7.01	–	–	–	–	–	–
g_{21}	16.93	7.55	0.02	0.02	0.62	5.95	17.61	1.09	0.43	0.38	8.91	7.10
g_{22}	0.12	0.03	10.70	11.61	0.02	9.95	–	–	–	–	–	–

that indicators such as the net domestic credit/GDP ratio and the current account balance/GDP ratio are quite significant for the classification of the countries.

Table 3.10 provides further insight into the differences between the two models, in terms of the significance attributed to the indicators that are common to both models. In particular, Table 3.10 presents the Kendall's τ rank correlation coefficient measuring the similarities between the rankings of the country risk indicators according to their significance in each utility function of models A and B. The results indicate significant similarities in three cases, involving the utility functions U_1 (at the 10% significance level), $U_{\sim 2}$ (at the 5% significance level), and $U_{\sim 3}$ (at the 1% significance level). Each of these three utility functions is employed at different stages of the hierarchical discrimination process employed in the MHDIS method. In particular, the utility function U_1 is used for the identification of high-income economies (first stage). The utility function $U_{\sim 2}$ is used at the second stage of the process (distinction between countries of the upper-middle income group and countries belonging in either the lower-middle income group or the low-income one). Finally, the utility function $U_{\sim 3}$ is used for the distinction between lower middle income economies and low-income economies. Therefore, at all three stages of the classification process, models A and B have some similarities in terms of the weight they assign to the considered country risk indicators.

Details on the performance of the two models in the classification of the countries in the sample are given in Table 3.11 for all the five years of the analysis. Overall, it is interesting to note that both models perform quite satisfactorily in the

Table 3.10 Similarities between the two country risk models of MHDIS (comparison of the rankings of the indicators according to their importance)

	U_1	$U_{\sim 1}$	U_2	$U_{\sim 2}$	U_3	$U_{\sim 3}$
Kendall's τ	0.438^{***}	-0.212	0.047	0.504^{**}	0.107	0.576^{*}

Notes: * Significant similarities at the 1% level, **significant similarities at the 5% level, *** significant similarities at the 10% level.

Table 3.11 Classification results of the MHDIS methods (in %)

		Model A (complete set of indicators)					Model B (reduced set of 12 indicators)				
Years		C_1	C_2	C_3	C_4	Overall	C_1	C_2	C_3	C_4	Overall
2000	C_1	100.00	0.00	0.00	0.00	99.43	100.00	0.00	0.00	0.00	94.25
	C_2	0.00	100.00	0.00	0.00		0.00	83.33	16.67	0.00	
	C_3	0.00	2.27	97.73	0.00		0.00	4.55	95.45	0.00	
	C_4	0.00	0.00	0.00	100.00		0.00	0.00	1.79	98.21	
1999	C_1	100.00	0.00	0.00	0.00	86.09	100.00	0.00	0.00	0.00	83.67
	C_2	0.00	83.33	13.33	3.33		13.33	60.00	20.00	6.67	
	C_3	0.00	20.45	68.18	11.36		0.00	9.09	81.82	9.09	
	C_4	0.00	1.79	5.36	92.86		0.00	0.00	7.14	92.86	
1998	C_1	100.00	0.00	0.00	0.00	82.99	100.00	0.00	0.00	0.00	83.43
	C_2	0.00	80.00	6.67	13.33		10.00	60.00	26.67	3.33	
	C_3	0.00	18.18	59.09	22.73		0.00	9.09	77.27	13.64	
	C_4	0.00	0.00	7.14	92.86		1.79	0.00	1.79	96.43	
1997	C_1	100.00	0.00	0.00	0.00	78.10	96.77	3.23	0.00	0.00	81.81
	C_2	0.00	76.67	13.33	10.00		0.00	66.67	26.67	6.67	
	C_3	0.00	18.18	50.00	31.82		0.00	11.36	72.73	15.91	
	C_4	0.00	3.57	10.71	85.71		0.00	0.00	8.93	91.07	
1996	C_1	96.77	3.23	0.00	0.00	78.08	100.00	0.00	0.00	0.00	81.52
	C_2	0.00	66.67	16.67	16.67		0.00	60.00	36.67	3.33	
	C_3	0.00	18.18	61.36	20.45		0.00	9.09	75.00	15.91	
	C_4	0.00	3.57	8.93	87.50		0.00	0.00	8.93	91.07	

classification of the countries in all the years of the analysis, even in 1996 (four years prior to the year 2000 upon which the development of the models was based), despite the complexity of the problem (four groups). The lower overall accuracy of model A is 78.08% for year 1996, whereas the overall accuracy of model B is consistently higher than 81% for all the years of the analysis.

A comparison of the two models indicates that model A performs better (in terms of overall accuracy) than model B in the two most recent years (2000 and 1999). Of course, considering that the data of the year 2000 are used as the reference set, the superiority of model A in this year should be attributed to the richer information that it considers (22 indicators in model A vs. 12 indicators in model B). For the remaining three years (1996–1998), model B performs consistently better than model A, with differences ranging between 0.44% (year 1998) and 3.71% (year 1997). Thus, it can be argued that model B that considers a compact set of information provides more robust results, throughout the period studied, compared with model A that considers all the available information on the 22 selected indicators. As far as the classification accuracies of the four groups are concerned, it should be noted that model A performs consistently better than model B in the classification of upper-middle income economies (group C_2), with differences ranging between 6.67% (year 1996) and 23.33% (year 1998). On the other hand, model B performs better than model A for groups C_3 and C_4 (lower-middle income economies and low-income economies, respectively).

Finally, it is worth noting that, overall, both models are more accurate for the two "extreme" groups C_1 and C_4 (high-income and low-income economies, respectively) compared with the two intermediate groups C_2 and C_3 (lower and upper-middle income economies, respectively). This should not be a surprise, as it is expected that high-income economies and low-income economies are easier to identify compared with intermediate cases such as lower-middle and upper-middle income economies.

3.3.2 Comparative Analysis with Other Approaches

In order to obtain a better understanding of the performance of the two MHDIS country risk assessment models, a comparison is performed with other well-known classification methodologies. The methodologies considered in this comparison involve a diversified set of approaches, including rough sets, artificial neural networks (ANN), the C5 algorithm, linear discriminant analysis (LDA), and logit analysis. These approaches are widely used in developing classification models and have been shown to be quite efficient in several domains. Therefore, the comparison of the MHDIS models with country risk models developed by these alternative approaches contributes to the clarification of the relative efficiency of the MHDIS method in developing country risk models as opposed to other established techniques.

The rough sets approach and the C5 algorithm implement the machine learning paradigm. The former leads to the development of classification models expressed in the form of "If ... then ..." decision rules. The implementation of the rough sets approach in this study is based on the use of the MODLEM algorithm for the development of the decision rules (Grzymala-Busse and Stefanowski, 2001), whereas the classification is performed on the basis of the paradigm of the LERS system (Grzymala-Busse, 1992). On the other hand, the C5 algorithm is a recent extension of C4.5, which is one of the most popular algorithms within the machine learning community for developing decision trees for classification (Quinlan, 1993). The third classification methodology considered in this comparison is a feed-forward ANN. The architecture of the ANN used in this study involves one hidden layer with five neurons, and the output layer consists of four neurons, one for each group of countries. Alternative architectures have also been explored, considering the number of hidden layers and the number of neurons in each of them, but the selected architecture provided the best results. The training of the ANN is based on the back-propagation approach using the scaled conjugate gradient algorithm (Moller, 1993). The final two methods used in this comparative analysis, the LDA and the logit analysis, are two widely used statistical and econometric techniques for developing classification models, with many applications in developing country risk assessment models (Saini and Bates, 1984).

Table 3.12 Comparison results (classification accuracies, in %)

		Complete set of indicators					Reduced set of 12 indicators				
Years	Methods	C_1	C_2	C_3	C_4	Overall	C_1	C_2	C_3	C_4	Overall
2000	MHDIS	100.0*	100.0*	97.7	100.0*	99.4	100.0*	83.3	95.5	98.2	94.3
	Rough sets	100.0*	100.0*	100.0*	100.0*	100.0*	100.0*	100.0*	100.0*	100.0*	100.0*
	ANN	100.0*	100.0*	95.5	96.4	98.0	96.8	90.0	90.9	92.9	92.6
	C5	100.0*	96.7	90.9	96.4	96.0	100.0*	96.7	93.2	98.2	97.0
	LDA	100.0*	80.0	65.9	83.9	82.5	96.8	76.7	59.1	76.8	77.3
	Logit	87.1	60.0	65.9	87.5	75.1	87.1	56.7	63.6	82.1	72.4
1999	MHDIS	100.0*	83.3*	68.2	92.9*	86.1	100.0*	60.0	81.8*	92.9*	83.7
	RS	96.8	73.3	88.64*	85.7	86.1	100.0*	70.0	79.6	83.9	83.4
	ANN	100.0*	76.7	77.3	87.5	85.4	100.0*	73.3*	81.8*	82.1	84.3*
	C5	100.0*	70.0	86.4	91.1	86.9*	100.0*	73.3*	72.7	91.1	84.3
	LDA	96.8	76.7	59.1	75.0	76.9	93.6	73.3*	56.8	82.1	76.5
	Logit	83.9	53.3	79.6	82.1	74.7	80.7	63.3	75.0	85.7	76.2
1998	MHDIS	100.0*	80.0*	59.1	92.9*	83.0*	100.0*	60.0	77.3*	96.4	83.4*
	Rough sets	96.8	56.7	68.2	89.3	77.7	96.8	56.7	65.9	78.6	74.5
	ANN	96.8	70.0	70.5	76.8	78.5	100.0*	76.7*	70.5	75.0	80.5
	C5	93.6	50.0	77.27*	89.3	77.5	100.0*	60.0	68.2	98.2*	81.6
	LDA	96.8	66.7	54.6	83.9	75.5	90.3	63.3	54.6	87.5	73.9
	Logit	80.7	33.3	59.1	83.9	64.3	67.7	53.3	68.2	87.5	69.2
1997	MHDIS	100.0*	76.7*	50.0	85.7*	78.1*	96.8	70.0*	72.7	91.1*	82.6*
	Rough sets	93.6	56.7	70.5	83.9	76.2	100.0*	63.3	63.6	83.9	77.7
	ANN	100.0*	56.7	75.0*	71.4	75.8	96.8	60.0	86.4*	73.2	79.1
	C5	96.8	50.0	75.0*	83.9	76.4	100.0*	53.3	63.6	85.7	75.7
	LDA	100.0*	63.3	68.2	73.2	76.2	100.0*	70.0*	59.1	76.8	76.5
	Logit	71.0	43.3	68.2	76.8	64.8	77.4	50.0	63.6	78.6	67.4
1996	MHDIS	96.8	66.7*	61.4	87.5*	78.1*	100.0*	60.0*	75.0*	91.1*	81.5*
	Rough sets	93.6	43.3	65.9	82.1	71.2	100.0*	46.7	68.2	80.4	73.8
	ANN	100.0*	40.0	72.73*	76.8	72.4	100.0*	36.7	75.0*	87.5	74.8
	C5	93.6	46.7	65.9	83.9	72.5	100.0*	46.7	59.1	87.5	73.3
	LDA	96.8	56.7	70.5	69.6	73.4	100.0*	60.0*	68.2	78.6	76.7
	Logit	71.0	43.3	59.1	82.1	63.9	71.0	40.0	75.0	80.4	66.6

Note: * Highest classification accuracies.

Table 3.12 presents details on the comparison of the two MHDIS models with the corresponding country risk models of the aforementioned methods. The presented results involved the classification accuracies for each group of countries, as well as the overall accuracy for each year of the analysis.

In terms of the overall accuracy, the consideration of either the complete set or the reduced set of 12 indicators leads to similar conclusions with regard to the performance of the two corresponding MHDIS models as opposed to the other approaches. In particular, in the two most recent years 2000 and 1999, the two MHDIS models provide similar results to the ones obtained by rough sets, ANN, and C5. For the other three years (1998, 1997, and 1996), however, the MHDIS models outperform all the methods considered in the comparison. For the complete set of indicators, the superiority (in terms of overall accuracy) of the MHDIS model A over the models of the other methods in the three years 1996–1998 ranges between 1.67% (compared with C5 in 1997) and 18.74% (compared with logit analysis in 1998), with an average of 6.64% (4.45% without considering logit analysis, which consistently provides lower accuracies). Similarly, for the reduced set of 12 indicators, the superiority of the MHDIS model B over the corresponding models of the other methods in the three years 1996–1998 ranges between 1.83% (compared with

C5 in 1998) and 15.23% (compared with logit analysis in 1997), with an average of 7.78% (6.02% without considering logit analysis, which consistently provides lower accuracies).

Further insight in the differences among the methods can be obtained considering the classification accuracies for the four individual groups. For the complete set of ratios, the model A of the MHDIS method performs consistently better (in all the years used for testing, i.e., 1996–1999) than the corresponding models of the other approaches in terms of its accuracy for groups C_2 (upper-middle income economies) and C_4 (low-income economies). For the high-income economies (group C_1), the differences among the MHDIS models and the other methods are small, whereas for the lower-middle income economies (group C_3), the MHDIS model provides consistently lower accuracies than the models of rough sets, ANN, and C5. For the reduced set of ratios, the overall superiority of the model B of the MHDIS method over the other models is attributed more to the robust results (across all the years) that it provides for groups C_2, C_3, and C_4 (for the high income group all methods except logit analysis provide similar results), rather than to its clear superiority for some of the groups (as in the case of the complete set of ratios).

3.4 The Study of Gjonca, Doumpos, Baourakis, and Zopounidis (2004)

The performance of the UTADIS and MHDIS methods and their applicability in country risk assessment are explored in this study.

3.4.1 Data Set Description

This application involves the assessment of the country risk for 125 countries from different geographical regions all over the world. The selection was based on the availability of the data for the countries. The data used are derived from the World Bank and refer to a five-year period (1995–1999). The countries in the sample are classified by the World Bank into four classes as follows (World Bank, 2001):

- High-income economies (class C_1): This group includes 28 countries, mostly European countries, as well as the United States, Australia, New Zealand, Canada, Japan, Hong Kong, Singapore, etc. These countries are considered as the world's top economies with a stable political and social development.
- Upper-middle income economies (class C_2): Twenty countries are included in this second group. They are from Europe, South and Eastern Asia, and Latin and South America. These countries cannot be considered as developed ones neither from the economic nor from the sociopolitical point of view. However, they do have some positive perspectives for future development.

- Lower-middle income economies (class C_3): The third group includes 37 countries from Eastern Europe, Asia, Africa, and South and Latin America. These countries are facing economic as well as social and political problems that make their future doubtful and uncertain.
- Low-income economies (class C_4): This final group consists of 40 countries, mostly from Africa and Asia, who face significant problems from all aspects.

For each country in the sample, a rich set of information was considered involving 38 indicators measuring different aspects of country risk including external trade, economic growth, inflation and exchange rates, balance of payments, tax policies, macroeconomic policies, structural transformations, etc.

Obviously, the incorporation of such a number of evaluation criteria would result in the development of a country risk assessment model with limited practical value, as the amount of information required to implement it would be costly and time-consuming to gather. To overcome this problem, the significance of the selected country risk indicators in discriminating the four groups of countries was tested for each year in the analysis through a one-way ANOVA. The corresponding results are presented in Table 3.13.

According to the ANOVA results, there are 14 indicators that are significant in discriminating the four classes of countries for all five years of the analysis. Also, there is one indicator that is significant in four years and three other indicators that are significant in three years. For the 14 indicators that were found significant for all five years, there were some high correlations. In particular, money and quasi money (M2)/GDP was found highly correlated with liquid liabilities (M3)/GDP (correlation coefficient 0.937), whereas total debt service/GDP was found highly correlated with total debt service/gross national income (correlation coefficient 0.978). To avoid such high correlations that may pose difficulties in model development and interpretation, it was decided to retain in the analysis only the indicators money and quasi money (M2)/GDP and total debt service/GDP, which have been found significant in previous studies on country risk analysis (Frank and Cline, 1971; Feder and Just, 1977; Balkan, 1982). It was also decided to consider in the analysis the ratio current account balance/GDP, which is found statistically significant in four years of the analysis (according to the ANOVA test), as well as the ratio gross international reserves/imports of goods and services, which is found statistically significant in three years of the analysis (according to the ANOVA test) and has been extensively used in previous studies on country risk evaluation (Zopounidis et al., 1998).

Therefore, on the basis of the above procedure considering the statistical significance of country risk indicators (ANOVA results), their correlations, and their relevance to country risk assessment as reported in the international literature, 14 indicators are finally selected to be included in the developed country risk assessment models (Table 3.14).

Table 3.13 ANOVA results for the significance of the country risk indicators

	Years				
Indicators	1999	1998	1997	1996	1995
Inflation, consumer prices (annual)	2.26	4.43*	1.39	5.08*	2.01
Inflation, GDP deflator (annual)	2.32	3.31*	1.45	1.02	2.19
Money & quasi money (M2)/GDP$^{(5)}$	4.21*	6.11*	5.91*	6.00*	5.76*
Money & quasi money annual growth$^{(3)}$	7.51*	2.40	4.08*	2.33	2.86*
Domestic credit provided by banking sector/GDP$^{(5)}$	4.90*	15.45*	31.80*	31.22*	27.35*
Liquid liabilities (M3)/GDP$^{(5)}$	4.59*	7.58*	7.00*	8.32*	7.85*
Current account balance/GDP$^{(4)}$	4.60*	7.64*	7.64*	2.64	4.99*
Exports of goods & services/GDP$^{(5)}$	3.35*	4.64*	4.64*	4.25*	4.24*
Exports of goods & services (annual growth)	0.74	1.50	1.50	0.61	0.93
External balance on goods & services/GDP$^{(5)}$	12.70*	9.46*	9.46*	6.56*	9.40*
Annual GDP growth	2.11	0.28	0.28	0.43	0.14
Annual GNP growth	3.16*	0.38	0.38	0.52	0.15
Annual GNP per capita growth	3.25*	0.24	0.24	1.09	0.75
Gross international reserves in months of imports	2.72*	2.42	2.42	1.94	2.93*
Imports of goods & services/GDP	1.10	1.47	1.47	1.06	0.92
Imports of goods & services (annual growth)	1.70	0.17	0.17	0.20	0.22
Gross national expenditure/GDP$^{(5)}$	7.74*	9.46*	3.37*	2.99*	3.63*
Industry, value added/GDP$^{(5)}$	11.95*	9.16*	7.53*	5.10*	5.78*
Industry, value added (annual growth)	1.66	0.54	0.54	0.83	0.33
Trade/GDP	1.90	2.61	2.08	2.31	2.23
Gross domestic savings/GDP$^{(5)}$	8.26*	16.23*	14.40*	9.97*	14.06*
Foreign direct investment/GDP	1.21	0.26	0.89	0.83	0.99
Foreign direct investment/Gross capital formation	0.59	0.15	0.91	1.15	1.29
Current account balance/Exp. of goods & services$^{(5)}$	5.84*	7.01*	6.79*	6.60*	5.02*
Gross int. reserves/GNI	2.60	2.28	1.53	2.52	3.08*
Gross int. reserves/GDP	0.73	2.13	1.51	2.52	3.12*
Gross int. reserves/Imports of goods & services$^{(3)}$	3.45*	2.69*	1.86	2.61	3.85*
Net int. reserves/Imports of goods & services	3.34*	2.63	1.87	2.21	3.10*
Growth in exports of goods & services/GNP growth	0.21	0.82	0.83	0.68	1.06
Growth in exports of goods & services/GDP growth	0.24	1.59	0.69	1.37	0.30
Growth in imports of goods & services/GNP growth	1.15	1.26	0.82	0.69	0.70
Growth in imports of goods & services/GDP growth	0.63	3.10	0.04	0.89	0.43
Gross capital formation/Total debt service$^{(5)}$	8.75*	5.00*	6.48*	3.95*	4.17*
Net capital account/Total debt service$^{(3)}$	5.12*	4.57*	5.71*	2.36	1.07
External debt/GDP$^{(5)}$	24.75*	18.33*	24.06*	21.64*	15.08*
Total debt service/GDP$^{(5)}$	20.98*	19.84*	18.41*	15.30*	4.16*
Total debt service/Exports of goods & services$^{(5)}$	14.91*	21.48*	19.73*	17.78*	8.31*
Total debt service/Gross national income$^{(5)}$	23.21*	19.07*	16.91*	12.42*	4.03*

Note: * Significant at the 5% level, (3) significant indicator in three years of the analysis, (4) significant indicator in four years of the analysis, (5) significant indicator in all years of the analysis.

3.4.2 Presentation of Results

In this study, the UTADIS and MHDIS methods were applied in the sample data to develop country risk models according to the classification provided by the World Bank. The most recent year is used as the training sample, and the previous years

Table 3.14 Selected country risk indicators

Abbreviations	Country risk indicators
M2/GDP	Money and quasi money (M2)/GDP
DC/GDP	Domestic credit provided by banking sector/GDP
CAB/GDP	Current account balance/GDP
EXP/GDP	Exports of goods and services/GDP
EBAL/GDP	External balance on goods and services/GDP
GNE/GDP	Gross national expenditure/GDP
IVA/GDP	Industry, value added/GDP
GDS/GDP	Gross domestic savings/GDP
CAB/EXP	Current account balance/Exp. of goods and services
GIR/IMP	Gross int. reserves/Imports of goods and services
GCF/TDS	Gross capital formation/Total debt service
ED/GDP	External debt/GDP
TDS/GDP	Total debt service/GDP
TDS/EXP	Total debt service/Exports of goods and services

are used to test the generalizing performance of the methods. The obtained results of the two methods are presented in this section.

Results of the UTADIS Method

The use of the UTADIS method on the data described earlier leads to the development of an additive utility model for the classification of the countries into the four predetermined classes. The model is developed on the basis of the most recent data available (year 1999); its ability to provide early warning signals on the risk level of the countries is tested by applying it to the previous years of the analysis (1995–1998) to examine whether the model's classifications coincide with the current state of the countries in the sample (predetermined classification). The classification results of the developed model are presented in Table 3.15. The elements $C_1 - C_1$, $C_2 - C_2$, $C_3 - C_3$, and $C_4 - C_4$ represent the classification accuracy for each of the four classes, and all the other elements correspond with classification errors.

With regard to the training sample (year 1999), the overall classification accuracy of UTADIS is 84.75%, whereas the accuracy of the model in the previous years gradually deteriorates. Generally, the model performs quite well in the most recent years (1999–1997), but its performance on the earliest years (1995–1996) drops below 73%. In terms of the individual error rates, the model performs excellent in identifying the countries of low risk (high-income economies, class C_1). Its ability to identify the high-risk countries (low-income economies, class C_4) is also satisfactory at least for the most recent years (1999–1997). On the other hand, the model's performance in the identification of the countries of medium risk (upper-middle and lower-middle income economies, classes C_2 and C_3) is moderate throughout all the years. However, it should be noted that throughout the five years of the analysis, there are no major errors of the form $C_1 \rightarrow C_4$ (classification of a high-income

Table 3.15 Classification results of UTADIS (in %)

Years	Original classification	Estimated classification C_1	C_2	C_3	C_4	Overall accuracy
1999	C_1	96.43	3.57	0.00	0.00	84.75
	C_2	5.00	85.00	10.00	0.00	
	C_3	5.41	13.51	67.57	13.51	
	C_4	0.00	5.00	5.00	90.00	
1998	C_1	100.00	0.00	0.00	0.00	80.69
	C_2	5.00	70.00	20.00	5.00	
	C_3	8.11	5.41	70.27	16.22	
	C_4	0.00	5.00	12.50	82.50	
1997	C_1	100.00	0.00	0.00	0.00	78.67
	C_2	5.00	70.00	20.00	5.00	
	C_3	5.41	16.22	62.16	16.22	
	C_4	0.00	5.00	12.50	82.50	
1996	C_1	100.00	0.00	0.00	0.00	72.99
	C_2	5.00	60.00	30.00	5.00	
	C_3	2.70	18.92	59.46	18.92	
	C_4	0.00	2.50	25.00	72.50	
1995	C_1	100.00	0.00	0.00	0.00	71.11
	C_2	15.00	55.00	25.00	5.00	
	C_3	8.11	16.22	59.46	16.22	
	C_4	0.00	5.00	25.00	70.00	

Table 3.16 Significance of country risk indicators in the UTADIS country risk model

Indicators	Weight (%)	Indicators	Weight (%)
M2/GDP	6.54	GDS/GDP	15.67
DC/GDP	2.91	CAB/EXP	0.28
CAB/GDP	7.97	GIR/IMP	13.52
EXP/GDP	10.78	GCF/TDS	0.10
EBAL/GDP	0.85	ED/GDP	20.05
GNE/GDP	5.82	TDS/GDP	15.33
IVA/GDP	0.07	TDS/EXP	0.09

economy as a low-income economy) or $C_4 \rightarrow C_1$ (classification of a low-income economy as a high-income economy).

Table 3.16 provides some details on the contribution (weight in the model) of each country risk indicator in the classification of the countries. Five indicators are found significant, having a cumulative weight in the model of more than 75%. These indicators are (1) external debt/GDP (ED/GDP, weight 20.05%), (2) gross domestic savings/GDP (GDS/GDP, weight 15.67%), (3) total debt service/GDP (TDS/GDP, weight 15.33%), (4) gross international reserves/imports of goods and services (GIR/IMP, weight 13.52%), and (5) exports of goods and services/GDP (EXP/GDP, weight 10.78%). Most of these indicators have been found significant in previous studies on country risk assessment (Feder and Uy, 1985; Balkan, 1992; Taffler and Abassi, 1984; Zopounidis et al., 1998).

Results of the MHDIS Method

The model development process and the classification of the countries in this four-class country risk problem are performed by the MHDIS method in three stages. In the first stage, a pair of additive utility functions $U_1(\mathbf{g})$ and $U_{\sim 1}(\mathbf{g})$ is developed to discriminate the low risk countries (high income economies) from all the other countries. In the second stage, a second pair of additive utility functions $U_2(\mathbf{g})$ and $U_{\sim 2}(\mathbf{g})$ is developed to discriminate the countries belonging in the upper-middle income economies class from all the countries belonging in the classes "lower middle income economies" and "low-income economies." Finally, in the third stage, the two utility functions $U_3(\mathbf{g})$ and $U_{\sim 3}(\mathbf{g})$ that are developed discriminate the countries belonging in the lower middle income economies class from the high-risk countries (low-income economies).

The classification results obtained from this hierarchical/sequential model development and classification approach of the MHDIS method are presented in Table 3.17. Similar to the UTADIS method, the most recent year (1999) is employed as the training sample for model development, whereas the previous years are used to test the ability of the model to provide early warning signals for the risk level of the countries. The most interesting finding on the performance of the model developed by MHDIS is that it provides quite robust results throughout all years of the analysis. The overall accuracy of the model is above 76% in all years, and furthermore its performance in the earliest years (1995–1996) is slightly better than the most recent years (1997–1998). Therefore, the model can be used to obtain quite reliable early warning signals for the level of country risk even for a period of up to five years. In terms of the individual error rates, similar to the UTADIS model, the country risk model of the MHDIS method performs quite well in identifying the low-risk and high-risk countries (classes C_1 and C_4), whereas its performance for the medium-risk countries (classes C_2 and C_3) is lower. Generally, throughout the five years of the analysis, the MHDIS model seems to be less accurate for the upper-middle income economies (class C_2) with accuracies ranging between 55% (years 1995 and 1998) up to 85% (training sample year 1999). It should also be noted that similar to the UTADIS model, in the case of the MHDIS method there are no major errors of the form $C_1 \rightarrow C_4$ (classification of a high-income economy as a low-income economy) or $C_4 \rightarrow C_1$ (classification of a low-income economy as a high-income economy).

Details on the contribution (weight) of the selected country risk indicators in the utility function developed by the MHDIS method for the classification of the countries are given in Table 3.18. Two indicators are found significant in all stages of the hierarchical discrimination process of the MHDIS model. These indicators include the gross international reserves/imports of goods and services ratio (GIR/IMP) and the external debt/GDP ratio (ED/GDP). Both these indicators were also found significant in the UTADIS model. The exports of goods and services/GDP ratio (EXP/GDP), which was found significant in the UTADIS model, has a weight of more than 10% in the utility functions $U_2(\mathbf{g})$, $U_{\sim 2}(\mathbf{g})$, $U_3(\mathbf{g})$, and $U_{\sim 3}(\mathbf{g})$ of the MHDIS model, thus indicating that it is significant in discriminating between

Table 3.17 Classification results of MHDIS (in %)

Years	Original classification	Estimated classification C_1	C_2	C_3	C_4	Overall accuracy
1999	C_1	100.00	0.00	0.00	0.00	95.57
	C_2	10.00	85.00	5.00	0.00	
	C_3	0.00	2.70	97.30	0.00	
	C_4	0.00	0.00	0.00	100.00	
1998	C_1	100.00	0.00	0.00	0.00	76.89
	C_2	5.00	55.00	25.00	15.00	
	C_3	2.70	10.81	67.57	18.92	
	C_4	0.00	2.50	12.50	85.00	
1997	C_1	100.00	0.00	0.00	0.00	76.89
	C_2	5.00	60.00	25.00	10.00	
	C_3	2.70	10.81	67.57	18.92	
	C_4	0.00	5.00	15.00	80.00	
1996	C_1	96.43	0.00	3.57	0.00	78.03
	C_2	5.00	65.00	25.00	5.00	
	C_3	0.00	8.11	75.68	16.22	
	C_4	0.00	2.50	22.50	75.00	
1995	C_1	96.43	0.00	3.57	0.00	78.08
	C_2	5.00	55.00	35.00	5.00	
	C_3	0.00	5.41	78.38	16.22	
	C_4	0.00	2.50	15.00	82.50	

Table 3.18 Significance of country risk indicators in the MHDIS country risk model (weights in %)

Indicators	Utility functions $U_1(\mathbf{g})$	$U_{\sim 1}(\mathbf{g})$	$U_2(\mathbf{g})$	$U_{\sim 2}(\mathbf{g})$	$U_3(\mathbf{g})$	$U_{\sim 3}(\mathbf{g})$
M2/GDP	2.42	3.34	2.07	2.29	14.65	11.75
DC/GDP	2.42	2.42	10.83	12.48	7.20	14.65
CAB/GDP	2.35	2.35	1.76	3.21	1.39	1.39
EXP/GDP	2.51	2.51	15.50	12.29	22.45	19.21
EBAL/GDP	15.55	11.73	12.44	2.05	1.67	1.67
GNE/GDP	5.39	18.68	12.98	27.98	1.65	1.65
IVA/GDP	7.20	2.27	1.98	1.98	1.63	1.63
GDS/GDP	11.11	2.53	2.05	2.59	1.67	1.67
CAB/EXP	2.20	8.45	1.67	1.67	1.32	1.61
GIR/IMP	18.20	10.01	9.73	15.91	15.51	9.12
GCF/TDS	2.00	2.00	2.00	2.00	1.65	4.29
ED/GDP	12.17	15.02	23.49	10.76	8.19	23.81
TDS/GDP	1.45	1.45	1.45	2.74	6.83	2.47
TDS/EXP	15.02	17.23	2.05	2.05	14.20	5.10

medium-risk and high-risk countries (classes C_2, C_3, and C_4). The gross domestic savings/GDP ratio (GDS/GDP), which was also found significant in the UTADIS model, in MHDIS it has a significant contribution only in the identification of the low-risk countries (high-income economies, class C_1), with a weight 11.11% in the corresponding utility function $U_1(\mathbf{g})$.

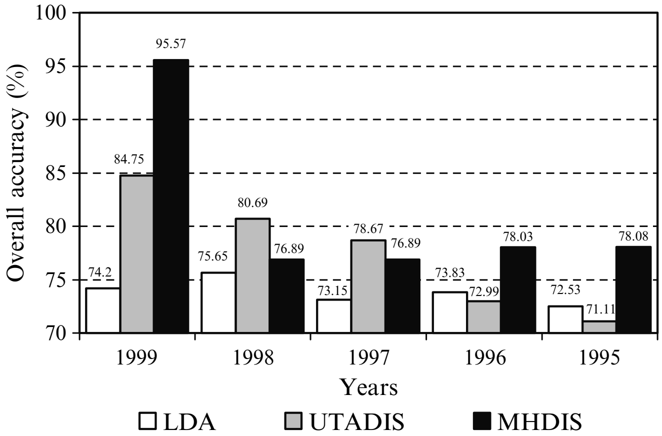

Fig. 3.3 Comparison of UTADIS and MHDIS to MDA

Comparison with Discriminant Analysis

Similar to the previous application, discriminant analysis (LDA) is also used in the analysis as a benchmark for the performance of the two MCDA models. The results of the comparison are outlined in Figure 3.3. The two MCDA models of UTADIS and MHDIS, generally, perform better than the LDA model. In the training sample, the MHDIS model provides considerably higher accuracy than the other two methods, but this is due to its larger number of degrees of freedom that provide the model with higher fitting ability. For the remaining years (1995–1998), which are used for testing purposes, the UTADIS model outperforms both the MHDIS model and the LDA model in the most recent years (1997–1998). In these two years, the differences between UTADIS and LDA in terms of the overall accuracy exceed 5% (in both years), whereas compared with the MHDIS model, the UTADIS model's accuracy is higher with differences ranging between 1.78% (1997) and 3.8% (1998). However, the performance of the UTADIS model in the two earliest years of the analysis (1995–1996) is lower compared with the other two methods. The differences in this case are small compared with the LDA model, but they are considerably higher compared with the MHDIS model (higher than 5% for both 1995 and 1996). On the other hand, the MHDIS model is consistently better than the LDA model in all years of the analysis with differences ranging between 1.24% (year 1998) up to 5.55% (year 1995).

3.5 The Study of Doumpos, Kosmidou, and Zopounidis (2004)

The performance of multicriteria analysis and non-parametric techniques and their applicability in country risk assessment are explored in this case study.

3.5.1 Data Set Description

The sample data used in this application is derived from the World Bank. From a total of 207 countries, only 161 countries, for which data were available for the examined period 1996–2000, were selected. Twenty-three indicators referring to inflation, exchange rates, the balance of payments, macroeconomics, etc., were selected.

According to the World Bank, the 161 countries are categorized into four classes based on their income level:

1. High-income economies (class C_1), including 31 countries (United States, Canada, New Zealand, Japan, European countries, etc.).
2. Upper-middle income economies (class C_2), including 30 countries, mostly countries from Latin America, as well as from Eastern Asia.
3. Lower-middle income economies (class C_3), including 44 countries, mostly from Asia, South and Latin America.
4. Low-income economies (class C_4), including 56 countries, mostly from Asia and Africa.

Similar to the previous application, in order to avoid the development of a country risk assessment model with limited practical value, the significance of the 23 indicators among the four groups of countries was tested through a one-way ANOVA. Table 3.19 presents the obtained results of the one-way ANOVA test.

The results indicate that only 16 indicators present statistically significant differences among the four groups at the 1% level. The criteria g_4, g_5, g_7, g_8, g_{15}, g_{17}, and g_{23} are not statistically significant and are excluded from the analysis. In order to have further analysis of the selected criteria, the correlations among the criteria are examined. Based on the correlation matrix of Table 3.20, the criteria that are excluded from the analysis due to high correlation values are g_{14}, g_{16}, g_{19}, and g_{21}. Table 3.21 presents the criteria that are finally selected for the analysis.

3.5.2 Presentation of Results

Results of Multicriteria Analysis

The UTADIS and MHDIS methods were applied to develop country risk models according to the classification provided by the World Bank.

Table 3.19 ANOVA results for the significane of the 23 indicators

	1996	1997	1998	1999	2000
g_1: Current account balance/GDP	0.013	0.000*	0.000*	0.001*	0.005*
g_2: Exports of goods and services/GDP	0.000*	0.001*	0.000*	0.001*	0.000*
g_3: Foreign direct investment/GDP	0.518	0.464	0.392	0.164	0.001*
g_4: GDP growth (annual %)	0.433	0.933	0.993	0.233	0.685
g_5: Growth of GDP per capita	0.393	0.542	0.516	0.194	0.050
g_6: Gross capital formation/GDP	0.165	0.007*	0.019	0.257	0.114
g_7: Gross fixed capital formation/GDP	0.235	0.023	0.076	0.380	0.241
g_8: Imports of goods and services/GDP	0.548	0.631	0.487	0.735	0.039
g_9: Inflation (per year %)	0.859	0.725	0.006*	0.201	0.594
g_{10}: Infant mortality rate (%)	0.000*	0.000*	0.000*	0.000*	0.000*
g_{11}: Short-term debt/Total external debt	0.000*	0.000*	0.000*	0.000*	0.000*
g_{12}: Total debt service/Exports of goods and services	0.000*	0.000*	0.000*	0.000*	0.000*
g_{13}: Total debt service/GNI	0.000*	0.000*	0.000*	0.000*	0.000*
g_{14}: Life expectancy at birth	0.000*	0.000*	0.000*	0.000*	0.000*
g_{15}: Consumer price index	0.859	0.935	0.619	0.622	0.662
g_{16}: GDP per capital	0.000*	0.000*	0.000*	0.000*	0.000*
g_{17}: Growth in export of goods and services	0.326	0.762	0.323	0.896	0.707
g_{18}: Net domestic credit/ GDP	0.000*	0.000*	0.000*	0.000*	0.000*
g_{19}: Total external debt/Exports of goods and services	0.000*	0.000*	0.000*	0.000*	0.000*
g_{20}: Total external debt/GDP	0.000*	0.000*	0.000*	0.000*	0.000*
g_{21}: Total debt service/Exports of goods and services	0.000*	0.000*	0.000*	0.000*	0.000*
g_{22}: Total debt service/Gross international reserves	0.318	0.439	0.503	0.527	0.000*
g_{23}: Gross international reserves in months of imports	0.291	0.489	0.669	0.313	0.163

Note: * Significance at the 1% level.

The classification results of the developed model based on the UTADIS method are presented in Table 3.22. The elements $C_1 - C_1$, $C_2 - C_2$, $C_3 - C_3$ and $C_4 - C_4$ represent the classification accuracy for each of the four classes, and all the other elements correspond with classification errors. Generally, the model performs well during the period 1996–2000 as its performance is more than 79%. Regarding the error rates, the model performs excellent in identifying the countries of low-risk (high-income economies, class C_1). It should also be noted that throughout the years there are no misclassifications of a high-income economy as a low-income one ($C_1 - C_4$), or of a low-income economy as a high-income one ($C_4 - C_1$).

Concerning the weights of the criteria, details are given in Table 3.23. The criterion infant mortality rate is most significant (30.58%) to the classification of countries, whereas the second significant one is the indicator current account balance/GDP with weight 16.88%.

The classification results of the MHDIS method are presented in Table 3.24. Similar to the UTADIS method, the performance of the model developed by MHDIS provides robust results throughout all the years of the analysis. The overall accuracy of the model is above 81%. Concerning the error rates, the country risk model of the MHDIS method performs quite well in identifying the low-risk and high-risk countries (C_1 and C_4, respectively). It should also be noted that similar to the UTADIS method, there are no major errors of the form $C_1 - C_4$ (classification of

Table 3.20 Correlation matrix of country risk indicators

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	g_{10}	g_{11}	g_{12}	g_{13}	g_{14}	g_{15}	g_{16}	g_{17}	g_{18}	g_{19}	g_{20}	g_{21}	g_{22}	g_{23}
g_1	1.00																						
g_2	0.20	1.00																					
g_3	−0.50	0.26	1.00																				
g_4	−0.13	0.13	0.19	1.00																			
g_5	−0.12	0.16	0.21	**0.96**	1.00																		
g_6	−0.44	0.32	0.51	0.37	0.38	1.00																	
g_7	−0.44	0.32	0.51	0.36	0.37	**0.98**	1.00																
g_8	−0.22	**0.83**	0.43	0.18	0.19	**0.55**	**0.55**	1.00															
g_9	0.17	0.08	0.07	−0.03	−0.03	−0.04	−0.03	0.01	1.00														
g_{10}	−0.17	−0.32	−0.06	0.09	−0.07	−0.15	−0.13	−0.16	0.08	1.00													
g_{11}	−0.09	0.01	0.01	0.03	0.04	0.11	0.12	0.01	0.01	−0.02	1.00												
g_{12}	−0.08	−0.33	−0.07	−0.13	−0.16	−0.12	−0.12	−0.29	0.03	0.29	0.23	1.00											
g_{13}	−0.09	0.10	0.08	−0.12	−0.13	0.10	0.09	0.11	0.21	0.15	0.28	**0.61**	1.00										
g_{14}	0.18	0.29	0.07	−0.06	0.09	0.15	0.13	0.14	−0.06	**−0.94**	−0.03	−0.28	−0.16	1.00									
g_{15}	−0.01	0.08	0.14	−0.01	−0.02	0.02	0.02	0.04	0.10	0.11	0.03	0.04	0.24	−0.07	1.00								
g_{16}	0.30	0.27	0.06	−0.03	0.05	−0.01	−0.01	0.06	−0.04	**−0.52**	−0.33	−0.41	−0.41	**0.56**	−0.02	1.00							
g_{17}	0.06	0.01	−0.02	0.33	0.33	0.09	0.09	0.02	0.07	−0.02	−0.02	−0.06	−0.26	0.02	**−0.56**	0.02	1.00						
g_{18}	0.11	0.21	0.07	−0.08	−0.02	0.09	0.10	0.11	−0.06	−0.46	−0.07	−0.20	−0.09	**0.53**	−0.05	**0.52**	0.03	1.00					
g_{19}	−0.24	−0.36	−0.08	−0.05	−0.13	−0.11	−0.10	−0.17	−0.04	**0.59**	0.05	0.40	0.12	**−0.53**	0.00	−0.33	0.04	−0.26	1.00				
g_{20}	−0.27	−0.12	0.01	−0.05	−0.13	0.07	0.08	0.09	0.01	0.51	0.08	0.38	0.43	−0.47	0.06	−0.41	0.02	−0.23	**0.76**	1.00			
g_{21}	0.06	−0.23	−0.07	−0.14	−0.16	−0.11	−0.12	−0.30	0.06	0.14	0.27	**0.90**	**0.68**	−0.14	0.05	−0.36	−0.08	−0.13	0.17	0.22	1.00		
g_{22}	−0.01	−0.02	−0.02	0.01	0.00	−0.01	0.00	−0.04	0.00	0.07	0.04	0.06	0.05	−0.08	0.00	−0.04	0.00	−0.05	0.03	0.06	0.08	1.00	
g_{23}	0.14	−0.17	−0.12	−0.06	−0.09	0.04	0.04	−0.15	−0.06	0.04	−0.03	0.16	−0.03	−0.08	−0.06	−0.11	0.01	−0.12	0.06	−0.01	0.17	−0.09	1.00

Table 3.21 Selected country risk indicators

g_1	Current account balance/GDP
g_2	Exports of goods and services/GDP
g_3	Foreign direct investment/GDP
g_4	Gross capital formation/GDP
g_5	Inflation (per year %)
g_6	Infant mortality rate (%)
g_7	Short-term debt/Total external debt
g_8	Total debt service/Exports of goods and services
g_9	Total debt service/GNI
g_{10}	Net domestic credit/GDP
g_{11}	Total external debt/GDP
g_{12}	Total debt service/Gross international reserves

Table 3.22 Results of the UTADIS classification model

		C_1	C_2	C_3	C_4	Overall accuracy
2000	C_1	96.77%	3.23%	0.00%	0.00%	83.96%
	C_2	3.33%	73.33%	23.33%	0.00%	
	C_3	0.00%	11.36%	81.82%	6.82%	
	C_4	0.00%	0.00%	16.07%	83.93%	
1999	C_1	96.77%	3.23%	0.00%	0.00%	81.43%
	C_2	0.00%	70.00%	26.67%	3.33%	
	C_3	0.00%	13.64%	75.00%	11.36%	
	C_4	0.00%	0.00%	16.07%	83.93%	
1998	C_1	96.77%	3.23%	0.00%	0.00%	79.76%
	C_2	0.00%	63.33%	33.33%	3.33%	
	C_3	0.00%	11.36%	75.00%	13.64%	
	C_4	0.00%	0.00%	16.07%	83.93%	
1997	C_1	96.77%	3.23%	0.00%	0.00%	80.57%
	C_2	0.00%	63.33%	36.67%	0.00%	
	C_3	0.00%	4.55%	81.82%	13.64%	
	C_4	0.00%	1.79%	17.86%	80.36%	
1996	C_1	96.77%	3.23%	0.00%	0.00%	79.23%
	C_2	0.00%	56.67%	40.00%	3.33%	
	C_3	0.00%	4.55%	79.55%	15.91%	
	C_4	0.00%	3.57%	12.50%	83.93%	

a high-income economy as a low-income economy) or $C_4 - C_1$ (classification of a low-income economy as a high-income economy).

The weights of the selected country risk indicators in the MHDIS method are presented in Table 3.25. The criterion Current account balance/GDP, which was found significant in the UTADIS method, has a weight of 16.55% in the utility function U_1, indicating that it is significant in discriminating high-risk countries (class C_1). The criterion Net domestic credit/GDP has a weight of 24.20% and 29.10% in the utility functions $U_{\sim 1}$ and $U_{\sim 3}$, respectively, indicating that it has a significant contribution in the identification of the high-risk and low-risk countries (classes C_1, C_4).

Table 3.23 Weights of the UTADIS classification model

Criteria	Weights
g_1: Current account balance/GDP	16.88%
g_2: Exports of goods and services/GDP	0.03%
g_3: Foreign direct investment/GDP	5.57%
g_4: Gross capital formation/GDP	3.69%
g_5: Inflation (annual %)	6.24%
g_6: Infant mortality rate	30.58%
g_7: Short-term debt/Total external debt	0.92%
g_8: Total debt service/Exports of goods and services	4.04%
g_9: Total debt service/Gross international reserves	2.00%
g_{10}: Net domestic credit/GDP	12.93%
g_{11}: Total external debt/GDP	10.13%
g_{12}: Total debt service/Gross international reserves	6.98%

Table 3.24 Results of the MHDIS classification model

		C_1	C_2	C_3	C_4	Overall accuracy
2000	C_1	100.00%	0.00%	0.00%	0.00%	94.25%
	C_2	0.00%	83.33%	16.67%	0.00%	
	C_3	0.00%	4.55%	95.45%	0.00%	
	C_4	0.00%	0.00%	1.79%	98.21%	
1999	C_1	100.00%	0.00%	0.00%	0.00%	83.67%
	C_2	13.33%	60.00%	20.00%	6.67%	
	C_3	0.00%	9.09%	81.82%	9.09%	
	C_4	0.00%	0.00%	7.14%	92.86%	
1998	C_1	100.00%	0.00%	0.00%	0.00%	83.43%
	C_2	10.00%	60.00%	26.67%	3.33%	
	C_3	0.00%	9.09%	77.27%	13.64%	
	C_4	1.79%	0.00%	1.79%	96.43%	
1997	C_1	96.77%	3.23%	0.00%	0.00%	81.81%
	C_2	0.00%	66.67%	26.67%	6.67%	
	C_3	0.00%	11.36%	72.73%	15.91%	
	C_4	0.00%	0.00%	8.93%	91.07%	
1996	C_1	100.00%	0.00%	0.00%	0.00%	81.52%
	C_2	0.00%	60.00%	36.67%	3.33%	
	C_3	0.00%	9.09%	75.00%	15.91%	
	C_4	0.00%	0.00%	8.93%	91.07%	

Results of Non-parametric Approaches

As it was referred to Chapter 2, the rough sets approach is developed based on the rough rules "if ... then ...". These rules are responsible for the classification of the countries. Table 3.26 presents the overall accuracy in the classification of the countries based on the rough sets approach. Its performance reaches an average accuracy of 81.87% throughout the four years of the analysis. The most recent year 2000 is used as the training sample with a classification accuracy of 100%. The classification accuracy is reduced in 1999 to 83.37%, whereas the lowest classification accuracy (73.80%) is presented at 1996.

Table 3.27 presents the analysis of the rough sets approach. Based on this analysis, 9 out of the 12 criteria are responsible for the classification of the countries.

Table 3.25 Weights of the MHDIS classification model (in %)

Criteria	U_1	$U_{\sim 1}$	U_2	$U_{\sim 2}$	U_3	$U_{\sim 3}$
g_1: Current account balance/GDP	16.55	0.78	6.57	28.06	22.36	2.19
g_2: Exports of goods & services/GDP	0.80	0.80	0.65	0.65	11.28	8.11
g_3: Foreign direct investment/GDP	10.88	0.73	0.59	9.69	0.76	3.11
g_4: Gross capital formation/GDP	17.44	0.79	12.58	0.64	3.49	0.50
g_5: Inflation (annual)	4.21	13.06	2.54	5.66	5.74	0.51
g_6: Infant mortality rate	0.77	47.95	26.25	19.84	11.33	20.49
g_7: Short-term debt/Total external debt	8.74	0.61	0.61	0.61	0.49	0.49
g_8: Total debt service/Exports of goods & services	15.38	2.67	7.75	1.79	0.49	11.71
g_9: Total debt service/Gross international reserves	0.60	0.60	6.93	0.60	12.03	0.49
g_{10}: Net domestic credit/GDP	6.60	24.20	33.90	4.29	6.99	29.10
g_{11}: Total external debt/GDP	0.42	6.73	1.20	27.80	16.12	16.21
g_{12}: Total debt service/Gross international reserves	17.61	1.09	0.43	0.38	8.91	7.10

Table 3.26 Classification results of the rough sets approach

		C_1	C_2	C_3	C_4	Overall accuracy
2000	C_1	100.00%	0.00%	0.00%	0.00%	100.00%
	C_2	0.00%	100.00%	0.00%	0.00%	
	C_3	0.00%	0.00%	100.00%	0.00%	
	C_4	0.00%	0.00%	0.00%	100.00%	
1999	C_1	100.00%	0.00%	0.00%	0.00%	83.37%
	C_2	0.00%	70.00%	30.00%	0.00%	
	C_3	0.00%	9.09%	79.55%	11.36%	
	C_4	0.00%	7.14%	8.93%	83.93%	
1998	C_1	96.77%	3.23%	0.00%	0.00%	74.48%
	C_2	0.00%	56.67%	36.67%	6.67%	
	C_3	0.00%	11.36%	65.91%	22.73%	
	C_4	0.00%	10.71%	10.71%	78.57%	
1997	C_1	100.00%	0.00%	0.00%	0.00%	77.72%
	C_2	0.00%	63.33%	33.33%	3.33%	
	C_3	0.00%	11.36%	63.64%	25.00%	
	C_4	0.00%	7.14%	8.93%	83.93%	
1996	C_1	100.00%	0.00%	0.00%	0.00%	73.80%
	C_2	0.00%	46.67%	46.67%	6.67%	
	C_3	0.00%	11.36%	68.18%	20.45%	
	C_4	0.00%	5.36%	14.29%	80.36%	

Regarding the first rule, the indicators Total debt service/Gross international reserves and Infant mortality rate are those that determine the classification of the countries. These indicators classify 29 countries to the high-income economy class, indicating the significance of this rule. Twenty four rules are developed in total. From these, only two cover the classifcation of the countries to the high-income economy (class C_1) using only the criteria g_6, g_{10}, and g_{12}. Seven rules are used for the classification of the countries to the income economy classes C_1 and C_2. Moreover, 9 criteria are used for the classification to the third income economy class and 8 criteria (g_7 is excluded) for the classification to the second income economy class. Finally, 8 rules and 8 criteria are used for the classification of a country to the last income economy class.

Table 3.27 Analysis of rough sets rules

1. If $g_{12} \geq -0.005 \wedge g_6 \geq -7.2$ then 1 [29]
2. If $g_{12} \geq -0.005 \wedge g_{10} \geq 0.475$ then 1 [2]
3. If $g_{12} < -0.005 \wedge g_6 \geq -19.425 \wedge g_{10} \geq 0.285 \wedge g_3 \geq 1.285$ then 2 [19]
4. If $-58.15 \leq g_6 < -19.425 \wedge g_{12} < -1.18$ then 2 [4]
5. If $g_9 < -10.38 \wedge g_2 < 31.455$ then 2 [2]
6. If $g_6 \geq -58.15 \wedge g_{10} < 0.135 \wedge g_1 \geq 13.17$ then 2 [1]
7. If $g_6 < -58.15 \wedge g_{11} \geq -0.49 \wedge 21.515 \leq g_2 < 36.82$ then 2 [1]
8. If $g_{12} \geq -0.005 \wedge g_6 < -7.2 \wedge g_{10} < 0.475 \wedge g_1 \geq -3.45$ then 2 [2]
9. If $g_{12} < -0.005 \wedge g_6 \geq -19.425 \wedge g_{10} < 0.135 \wedge g_1 \geq 4.015$ then 2 [1]
10. If $-58.15 \leq g_6 < -19.425 \wedge g_{12} \geq -1.18 \wedge g_{10} \geq 0.135 \wedge g_9 \geq -10.38 \wedge g_2 < 50.345$ then 3 [24]
11. If $g_6 \geq -19.425 \wedge 0.135 \leq g_{10} < 0.285 \wedge g_7 < -3.105 \wedge g_9 \geq -11.03$ then 3 [7]
12. If $g_6 < -19.425 \wedge g_{10} \geq 0.135 \wedge g_9 < -10.38 \wedge g_2 \geq 49.505$ then 3 [3]
13. If $g_6 < -19.425 \wedge g_{10} \geq 0.135 \wedge g_2 \geq 55.055$ then 3 [3]
14. If $g_6 < -58.15 \wedge g_{11} \geq -0.49 \wedge g_2 \geq 36.82$ then 3 [4]
15. If $g_{12} < -0.005 \wedge g_6 \geq -19.425 \wedge g_3 < 1.285$ then 3 [2]
16. If $-58.15 \leq g_6 < -19.425 \wedge g_{12} \geq -1.18 \wedge 5.88 \leq g_1 < 13.17$ then 3 [1]
17. If $g_6 < -58.15 \wedge g_{11} < -0.49$ then 4 [39]
18. If $-58.15 \leq g_6 < -19.425 \wedge g_{10} < 0.135 \wedge g_1 < 5.88$ then 4 [6]
19. If $g_6 \geq -19.425 \wedge 0.135 \leq g_{10} < 0.285 \wedge g_7 \geq -3.105$ then 4 [2]
20. If $g_6 < -19.425 \wedge g_{10} \geq 0.135 \wedge g_9 < -10.38 \wedge 31.455 \leq g_2 < 49.505$ then 4 [2]
21. If $g_{12} < -0.005 \wedge g_6 \geq -19.425 \wedge g_{10} < 0.135 \wedge g_1 < 4.015$ then 4 [2]
22. If $g_6 < -58.15 \wedge g_2 < 21.515$ then 4 [3]
23. If $-58.15 \leq g_6 < -19.425 \wedge 50.345 \leq g_2 < 55.055$ then 4 [1]
24. If $0.135 \leq g_6 < 0.285 \wedge g_9 < -11.03$ then 4 [1]

According to the neural networks approach, the average overall accuracy is 82.27% and overpasses that of UTADIS and the rough sets approach. Table 3.28 presents the results of the neural networks. Similar to the previous methodologies, the highest classification accuracy (92.64%) is realized at 2000.

Results of Statistical Approaches

Apart from the multricriteria analysis and the non-parametric approaches, the discriminant and logistic analyses are also used. Tables 3.29 and 3.30 present the results of the discriminant and logistic analyses respectively, indicating that the overall accuracy in the classification of the countries throughout all the years is above 73% for the discriminant analysis and above 66% for the logistic analysis.

Figure 3.4 summarizes the comparative results (overall accuracies) of all the methods used in this analysis. It is obvious that the MHDIS method presents the highest accuracy (82.61% throughout the four years of the analysis and then follows the UTADIS method with an average accuracy of 80.25%). The neural networks have an accuracy rate of 79.68%, whereas the rough sets approach present accuracy of 77.34%. Based on the above results, the two MCDA models of UTADIS and MHDIS perform better than the discriminant and logistic analyses.

Table 3.28 Classification results of neural networks

		C_1	C_2	C_3	C_4	Overall accuracy
2000	C_1	96.77%	0.00%	3.23%	0.00%	92.64%
	C_2	3.33%	90.00%	6.67%	0.00%	
	C_3	0.00%	2.27%	90.91%	6.82%	
	C_4	0.00%	0.00%	7.14%	92.86%	
1999	C_1	100.00%	0.00%	0.00%	0.00%	84.32%
	C_2	6.67%	73.33%	20.00%	0.00%	
	C_3	0.00%	15.91%	81.82%	2.27%	
	C_4	1.79%	1.79%	14.29%	82.14%	
1998	C_1	100.00%	0.00%	0.00%	0.00%	80.53%
	C_2	6.67%	76.67%	13.33%	3.33%	
	C_3	0.00%	27.27%	70.45%	2.27%	
	C_4	1.79%	5.36%	17.86%	75.00%	
1997	C_1	96.77%	3.23%	0.00%	0.00%	79.09%
	C_2	3.33%	60.00%	33.33%	3.33%	
	C_3	0.00%	9.09%	86.36%	4.55%	
	C_4	0.00%	3.57%	23.21%	73.21%	
1996	C_1	100.00%	0.00%	0.00%	0.00%	74.79%
	C_2	6.67%	36.67%	50.00%	6.67%	
	C_3	2.27%	6.82%	75.00%	15.91%	
	C_4	0.00%	1.79%	10.71%	87.50%	

Table 3.29 Discriminant analysis results

		C_1	C_2	C_3	C_4	Overall accuracy
2000	C_1	96.77%	0.00%	3.23%	0.00%	77.33%
	C_2	10.00%	76.67%	10.00%	3.33%	
	C_3	4.55%	25.00%	59.09%	11.36%	
	C_4	0.00%	1.79%	21.43%	76.79%	
1999	C_1	93.55%	0.00%	6.45%	0.00%	76.46%
	C_2	10.00%	73.33%	10.00%	6.67%	
	C_3	4.55%	25.00%	56.82%	13.64%	
	C_4	0.00%	1.79%	16.07%	82.14%	
1998	C_1	90.32%	0.00%	6.45%	3.23%	73.93%
	C_2	6.67%	63.33%	23.33%	6.67%	
	C_3	0.00%	31.82%	54.55%	13.64%	
	C_4	0.00%	7.14%	5.36%	87.50%	
1997	C_1	100.00%	0.00%	0.00%	0.00%	76.47%
	C_2	6.67%	70.00%	20.00%	3.33%	
	C_3	2.27%	27.27%	59.09%	11.36%	
	C_4	0.00%	5.36%	17.86%	76.79%	
1996	C_1	100.00%	0.00%	0.00%	0.00%	76.69%
	C_2	3.33%	60.00%	33.33%	3.33%	
	C_3	0.00%	20.45%	68.18%	11.36%	
	C_4	3.57%	1.79%	16.07%	78.57%	

Table 3.30 Logistic regression results

		C_1	C_2	C_3	C_4	Overall accuracy
2000	C_1	87.10%	9.68%	3.23%	0.00%	72.39%
	C_2	13.33%	56.67%	16.67%	13.33%	
	C_3	0.00%	18.18%	63.64%	18.18%	
	C_4	0.00%	0.00%	17.86%	82.14%	
1999	C_1	80.65%	12.90%	6.45%	0.00%	76.17%
	C_2	6.67%	63.33%	16.67%	13.33%	
	C_3	0.00%	9.09%	75.00%	15.91%	
	C_4	0.00%	0.00%	14.29%	85.71%	
1998	C_1	67.74%	22.58%	6.45%	3.23%	69.19%
	C_2	3.33%	53.33%	30.00%	13.33%	
	C_3	0.00%	13.64%	68.18%	18.18%	
	C_4	0.00%	0.00%	12.50%	87.50%	
1997	C_1	77.42%	22.58%	0.00%	0.00%	67.41%
	C_2	3.33%	50.00%	33.33%	13.33%	
	C_3	0.00%	13.64%	63.64%	22.73%	
	C_4	0.00%	0.00%	21.43%	78.57%	
1996	C_1	70.97%	29.03%	0.00%	0.00%	66.58%
	C_2	10.00%	40.00%	30.00%	20.00%	
	C_3	0.00%	9.09%	75.00%	15.91%	
	C_4	0.00%	0.00%	19.64%	80.36%	

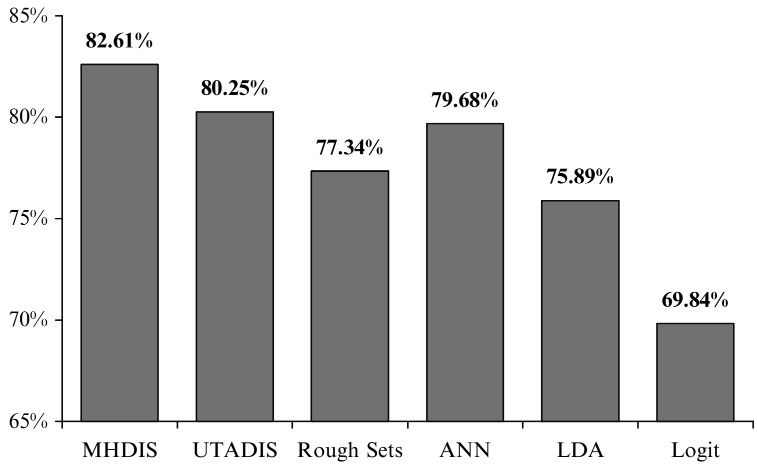

Fig. 3.4 Average accuracy rate of the classification models

Chapter 4
Conclusions

Country risk assessment is a multidimensional problem that is of major interest to policymakers, to managers of international lending institutions, to multinational firms, and to investors.

The chapters of this book presented the problem of country risk assessment and demonstrated the use of multicriteria methods, statistical and econometric techniques, as well as non-parametric techniques in the problem. The presentation of the methods as well as their applications indicated the advantages of the use of the methods. Moreover, the use of these approaches provides a guide for country risk assessment by practitioners and researchers.

More specifically, the country risk problem was studied in the first application as a ranking and sorting problem. In both cases, the obtained results are very satisfactory as the obtained country risk models are consistent with the preferences and the decision policy of the World Bank and Euromoney. The use of the five methods (UTADIS, I, II, III and UTASTAR) illustrated their ability in deriving flexible decision models taking into account the preferences of the decision makers. The decision maker plays a significant role in the decision process by interacting with the methods to take decisions in real time. The MHDIS method, which was used in the second and third application, constitutes a computational tractable procedure to develop additive utility models that can be used for classification purposes in financial risk assessment. The application of the MHDIS method in the assessment of country risk has demonstrated its efficiency in the analysis of complex real-world decision problems regarding financial risk assessment. The comparison with other well-known classification approaches demonstrated the efficiency of the proposed MCDA approach in addressing the country risk assessment problem. Both the UTADIS and the MHDIS methods led to the development of country risk classification models that provided high accuracies in classifying the countries for a time period of up to five years, as presented in the fourth application. Finally, the use of multicriteria analysis and non-parametric approaches in the assessment of country risk is presented in the last application. These methods are free of restrictive statistical assumptions, they are capable of incorporating in the decision process qualitative social and political factors, and they can be easily adapted to the changes in the decision environment.

K. Kosmidou et al. (eds.) *Country Risk Evaluation,* DOI: 10.1007/978-0-387-76680-5_4,

Such approaches provide decision makers (financial, credit, stock market analysts, investors, etc.) with a valuable tool to perform real-time evaluations on the financial risks of the considered alternatives. These methods take into account the knowledge and the preferences of the user in constructing the models. This fact allows the user to make any corrections or modifications if changes in the economic environment or other factors take place. Moreover, these methods accept and easily manage both quantitative and qualitative characteristics of the countries, enabling the decision maker to employ them in the country risk assessment.

Future research is required using a broader set of data related to the social and political aspects of country risk. This could contribute positively in facilitating an integrated analysis of country risk assessment. It would be interesting to explore the combination of different model development approaches for the development of more accurate and efficient country risk models, providing more information to analysts in the study of country risk.

References

1. Abassi, B. and Taffler, R.J. (1982) "Country risk: A model of economic performance related to debt servicing capacity," Working paper 36, City University Business School, London.
2. Agmon, T. and Deitrich, J.K. (1983) "International lending and income redistribution: An alternative view of country risk," *Journal of Banking and Finance*, 7, 483–495.
3. Altman, E. I., Avery, R., Eisenbeis, R., and Stinkey, J. (1981) *Application of Classification Techniques in Business, Banking and Finance*, JAI Press, Greenwich, CT.
4. Archer, N.P. and Wang, S. (1993) "Application of the back propagation neural networks algorithm with monotonicity constraints for two-group classification problems," *Decision Sciences*, 24, 60–75.
5. Balkan E.M. (1992) "Political instability, country risk and probability of default," *Applied Economics*, 24, 999–1008.
6. Bastian, A. (2000) "Identifying fuzzy models utilizing genetic programming," *Fuzzy Sets and Systems*, 113, 333–350.
7. Benayoun, R., De Montgolfier, J., Tergny, J. and Larichev, O. (1971) "Linear programming with multiple objective functions: Stem method (STEM)," *Mathematical Programming*, 1(3), 366–375.
8. Bird G. (1986) "New approaches to country risk," *Lloyds Bank Review*, 162, 1–16.
9. Breiman, L., Friedman, J.H., Olsen, R.A. and Stone, C.J. (1984) *Classification and Regression Trees*, Wadsworth International Group, Belmont, CA.
10. Brewer, T.L. and Rivoli, P. (1990) "Politics and perceived country creditworthiness in international banking," *Journal of Money, Credit and Banking*, 22(3), 357–369.
11. Burton, F.N. and Inoue, H. (1983) "Country risk evaluation methods: A survey of systems in use," *The Banker*, 133, 41–43.
12. Burton, F.N. and Inoue, H. (1987) "A country risk appraisal model of foreign asset expropriation in developing countries," *Applied Economics*, 19, 1009–1048.
13. Calverley, J. (1990) *Country Risk Analysis (2nd Edition)*, Butterworth and Co Publishers Ltd., Scotland.
14. Cantor, R. and Packer, E. (1996) "Determinants and impact of sovereign credit ratings," *Economic Policy Review*, 2(2), 37–53.
15. Citron, J.T. and Nickelsburg, G. (1987) "Country risk and political instability," *Journal of Development Economics*, 25, 385–392.
16. Charnes, A., Cooper, W.W. and Ferguson, R.O. (1955) "Optimal estimation of executive compensation by linear programming," *Management Science*, 2, 138–151.
17. Chmielewski, M.R. and Grzymala-Busse, J.W. (1996), "Global discretization of continuous attributes as preprocessing for machine learning," *International Journal of Approximate Reasoning*, 15, 319–331.
18. Clark, P. and Niblett, T. (1989) "The CN2 induction algorithm," *Machine Learning*, 3, 261–283.

19. Cook, W.D. and Hebner, J.H. (1993) "A multicriteria approach to country risk evaluation: With an example employing Japanese Data," *International Review of Economics and Finance*, 2(4), 327–348.
20. Cook, W.D. and Kress, M. (1991) "A multiple criteria decision model with ordinal preference data," *European Journal of Operational Research*, 54, 191–198.
21. Cooper, J.C.B. (1987) "Country creditworthiness: The use of cluster analysis and discriminant analysis," Glasgow College of Technology, Department of Economics, Discussion Papers, no. 6.
22. Cosset, J.C. and Roy, J. (1988) "Expert judgments of political riskiness: An alternative approach," Document de Travail 88–12, Université Laval, Quebec, Canada.
23. Cosset, J.C. and Roy, J. (1989) "The determinants of country risk ratings," Document de Travail 89–43, Université Laval, Quebec, Canada.
24. Cosset, J.C. and Roy, J. (1991) "The determinants of country risk ratings," *Journal of International Business Studies*, First Quarter, 135–142.
25. Cosset, J.C., Siskos Y. and Zopounidis, C. (1992) "Evaluating country risk: A decision support approach," *Global Finance Journal*, 3(1), 79–95.
26. Devaud, J.M., Groussaud, G. and Jacquet-Lagrèze, E. (1980) "UTADIS: Une méthode de construction de fonctions d'utilité additives rendant compte de jugements globaux," European Working Group on Multicriteria Decision Aid, Bochum.
27. Dhonte, P. (1975) "Describing external debt situations: A roll-over approach," IMF Staff Papers 22, 159–186.
28. Dillon, W.R. and Goldstein, M. (1978) "On the performance of some multinomial classification rules," *Journal of the American Statistical Association*, 73, 305–313.
29. Doumpos, M. and Zopounidis, C. (1998) "The use of the preference disaggregation analysis in the assessment of financial risks," *Fuzzy Economic Review*, 3/1, 39–57.
30. Doumpos, M. and Zopounidis, C. (2000) "Assessing financial risks using a multicriteria sorting procedure: The case of country risk assessment," *Omega*, 29 (1), 97–109.
31. Doumpos, M. and Zopounidis, C. (2002) *Multicriteria Decision Aid Classification Methods*, Kluwer Academic Publishers, Dordrecht.
32. Doumpos, M. and Zopounidis, C. (2002) "On the use of a multicriteria hierarchical discrimination approach for country risk assessment," *Journal of Multi-criteria Decision Analysis*, 11 (4–5), 279–289.
33. Doumpos, M., Kosmidou, K. and Zopounidis, C. (2004) *Analysis and Evaluation of Country Risk: Theoretical and Empirical Approach*, Klidarithmos, Athens (in Greek).
34. Dubois, D. and Prade, H. (1979) "Decision-making under fuzziness." In: *Advances in Fuzzy Set Theory and Applications* (Gupta, M.M., Ragade, R.K. and Yager, R.R., Eds) North-Holland, Amsterdam, 279–302.
35. Duda, R.O. and Hart, P.E. (1978) *Pattern Classification and Scene Analysis*, John Wiley and Sons, New York.
36. Erb, C.B. Harvey C.R. and Viskanta T.E. (1996) "Political risk, economic risk and financial risk," *Financial Analysts Journal*, 52(6), 29–46.
37. Fayyad, U.M. and Irani, K.B. (1992) "On the handling of continuous-valued attributes in decision tree generation," *Machine Learning*, 8, 87–102
38. Feder, G. and Just, R. (1977) "A study of debt servicing capacity applying logit analysis," *Journal of Development Economics*, 4, 25–38.
39. Feder, G. and Uy, L.V. (1985) "The determinants of international creditworthiness and their policy implications," *Journal of Policy Modeling*, 7(1), 133–156.
40. Fisher, R.A. (1936) "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, 7, 179–188.
41. Fodor, J. and Roubens, M. (1994) *Fuzzy Preference Modelling and Multicriteria Decision Support*, Kluwer Academic Publishers, Dordrecht.
42. Frank, C.R. and Cline, R. (1971) "Measurement of debt servicing capacity: An application of discriminant analysis," *Journal of International Economics*, 1, 327–344.

43. Gelfand, S., Ravishankar, C. and Delp, E. (1991) "An iterative growing and pruning algorithm for classification tree design," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(2), 163–174.
44. Gjonca, E., Doumpos, M., Baourakis, G. and Zopounidis, C. (2004) "Assessing country risk using multicriteria classification approaches." In: *Supply Chain and Finance* (Pardalos, P.M., Migdalas, A. and Baourakis, G., Eds), World Scientific Publishing, Singapore, 50–67.
45. Goldstein M., Kaminsky G.L. and Reinhart C.M. (2000) "Assessing financial vulnerability: An early warning system for emerging markets," Institute for International Economics, Washington, DC.
46. Grabisch, M. (1995) "Fuzzy integrals in multicriteria decision making," *Fuzzy Sets and Systems*, 69, 279–298.
47. Grabisch, M. (1996) "The application of fuzzy integrals in multicriteria decision making," *European Journal of Operational Research*, 89, 445–456.
48. Greco, S., Matarazzo, B. and Slowinski, R. (1997) "Rough set approach to multi-attribute choice and ranking problems." In: *Multiple Criteria Decision Making* (Fandel, G. and Gal, T., Eds), Springer-Verlag, Berlin, 318–329.
49. Greco, S., Matarazzo, B. and Slowinski, R. (1999a) "The use of rough sets and fuzzy sets in MCDM." In: *Advances in Multiple Criteria Decision Making* (Gal, T., Hanne, T. and Stewart, T., Eds.), Kluwer Academic Publishers, Dordrecht, 14.1–14.59.
50. Greco, S., Matarazzo, B., Slowinski, R. and Zanakis, S. (1999b) "Rough set analysis of information tables with missing values." In: *Integrating Technology & Human Decisions: Bridging into the 21st Century*, Vol. II, Proceedings of the 5th International Meeting of the Decision Sciences Institute (Despotis, D. and Zopounidis, C., Eds.), New Technologies Editions, Athens, 1359–1362.
51. Greco, S., Matarazzo, B. and Slowinski, R. (2000a) "Extension of the rough set approach to multicriteria decision support," *INFOR*, 38(3), 161–196.
52. Greco, S., Matarazzo, B. and Slowinski, R. (2000b) "Dealing with missing values in rough set analysis of multi-attribute and multi-criteria decision problems." In: *Decision Making: Recent Developments and Worldwide Applications* (Zanakis, S.H., Doukidis, G. and Zopounidis, C., Eds.), Kluwer Academic Publishers, Dordrecht, 295–316.
53. Greco, S., Matarazzo, B. and Slowinski, R. (2001), "Conjoint measurement and rough sets approach for multicriteria sorting problems in presence of ordinal data." In: *AMCDA-Aide Multicritère à la decision (Multiple Criteria Decision Aiding)* (Colorni, A., Paruccini, M. and Roy, B., Eds.), EUR Report, Joint Research Centre, The European Commission, Ispra, 117–144.
54. Greco, S., Matarazzo, B. and Slowinski, R. (2002) "Rough sets methodology for sorting problems in presence of multiple attributes and criteria," *European Journal of Operational Research*, 138, 247–259.
55. Grinols, E. (1976) "International debt rescheduling and discrimination using financial variables," Manuscript, U.S. Treasury Department, Washington, DC.
56. Grzymala-Busse, J.W. (1992) "LERS: system for learning from examples based on rough sets." In: *Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory* (Slowinski R., Ed.), Kluwer Academic Publishers, Dordrecht, 3–18.
57. Grzymala-Busse, J.W and Stefanowski, J. (2001) "Three discretization methods for rule induction," *International Journal of Intelligent Systems*, 26: 29–38.
58. Haque, N., Kumar, M.S. Mark N. and Mathieson, D.J. (1996) "The economic content of indicators of developing country creditworthiness," IMF Staff Papers, 96(9).
59. He, J., Liu, X., Shi, Y., Xu, W. and Yan, N. (2004) "Classification of credit cardholder behavior by using fuzzy linear programming," *International Journal of Information Technology and Decision Making*, 3, 633–650.
60. Heffernan, S.A. (1985) "Country risk analysis: The demand and supply of sovereign loans," *Journal of International Money and Finance*, 4, 389–413.
61. Herring, R. (1983) *Managing International Risk*, Cambridge University Press, Cambridge.
62. Hung, M.S. and Denton, J.W. (1993) "Training neural networks with the GRG2 nonlinear optimizer," *European Journal of Operational Research*, 69, 83–91.

63. Ishibuchi, H., Nozaki, K. and Tanaka, H. (1992) “Distributed representation of fuzzy rules and its application to pattern classification,” *Fuzzy Sets and Systems*, 52, 21–32.
64. Ishibuchi, H., Nozaki, K. and Tanaka, H. (1993) “Efficient fuzzy partition of pattern space for classification problems,” *Fuzzy Sets and Systems*, 59, 295–304.
65. Inuiguchi, M., Tanino, T. and Sakawa, M. (2000) “Membership function elicitation in possibilistic programming problems,” *Fuzzy Sets and Systems*, 111, 29–45.
66. Jacquet-Lagrèze, E. (1995) “An application of the UTA for the evaluation of R & D projects.” In: *Advances in Multicriteria Analysis* (Pardalos, P.M., Siskos, Y. and Zopounidis, C., Eds.), Kluwer Academic Publishers, Dordrecht, 203–211.
67. Jacquet-Lagrèze E. and Siskos Y. (1982) “Assessing a set of additive utility functions for multicriteria decision making: The UTA method,” *European Journal of Operational Research*, 10, 151–164.
68. Kaminsky G. and Schmukler S.L. (2002) “Emerging market instability: Do sovereign ratings affect country risk and stock returns?,” *World Bank Economic Review*, 16, 171–195.
69. Kobrin, S. (1986) “Assessing political risk overseas.” In: *Multinational Enterprise in Transition: Selected Readings and Essays* (Grub, P.D., Ed.), Darwin, London, 425–439.
70. Kodratoff, Y. and Michalski, R.S. (1990) *Machine Learning: An Artificial Intelligence Approach*, Volume III, Morgan Kaufmann Publishers, Los Altos, CA.
71. Kosko, B. (1992) *Neural Networks and Fuzzy Systems*, Prentice-Hall, Englewood Cliffs, NJ.
72. Kou, G., Liu, X., Peng, Y., Shi, Y., Wise, M. and Xu, W. (2003) “Multiple Criteria Linear Programming to Data Mining: Models, Algorithm Designs and Software Developments,” *Optimization Methods and Software*, 18, 453–473.
73. Kou, G., Peng, Y., Shi, Y., Wise, M. and Xu, W. (2005) “Discovering credit cardholders’ behavior by multiple criteria linear programming,” *Annals of Operations Research*, 135, 261–274.
74. Krzanowski, W.J. (1975) “Discrimination and classification using both binary and continuous variables,” *Journal of the American Statistical Association*, 70, 782–790.
75. Krzanowski, W.J. (1977) “The performance of Fisher’s linear discriminant function under nonoptimal conditions,” *Technometrics*, 19, 191–200.
76. Kugel, Y. (1973) “A decisional model for the multinational firm,” *Management International Review*, 13, 34–42.
77. Lanchenbruch, P.A., Snuringer, C. and Revo, L.T. (1973) “Robustness of the linear and quadratic discriminant function to certain types of non-normality,” *Communications in Statistics*, 1, 39–56.
78. Langholz, G., Kandel, A., Schneider, M. and Chew, G. (1996) *Fuzzy Expert System Tools*, John Wiley and Sons, New York.
79. Lee, S.H. (1993) “Relative importance of political instability and economic variables on perceived country creditworthiness,” *Journal of International Business Studies*, 24(4), 801–812.
80. Lindstone, H.A. and Turoff M. (1975) *The Delphi Method: Techniques and Applications*, Addison-Wesley, Reading, MA.
81. Lootsma, F.A. (1997) *Fuzzy Logic for Planning and Decision Making*, Kluwer Academic Publishers, Dordrecht.
82. Mayo A.L. and Barrett A.G. (1978) “An early warning model for assessing developing country risk.” In: *Financing and Risk in Developing Countries* (Goodman, S.H., Ed.), Praeger Publishers, New York.
83. Michalski, R.S. (1969) “On the quasi-minimal solution of the general covering problem,” Proceedings of the 5th International Federation on Automatic Control, Vol. 27, 109–129.
84. Mienko, R., Stefanowski, J., Toumi, K. and Vanderpooten, D. (1996) “Discovery oriented induction of decision rules,” Cahier du LAMSADE no. 141, Université de Paris Dauphine, Paris.
85. Moller, MF. (1993) “A scaled conjugate gradient algorithm for fast supervised learning,” *Neural Networks*, 6, 525–533.
86. Moore, D.H. (1973) “Evaluation of five discriminant procedures for binary variables,” *Journal* of the American Statistical Association, 68, 399–404.

87. Mondt, K. and Despontin, M. (1986) "Evaluation of country risk using multicriteria analysis," Technical Report, Vrije Universiteiet Brussels, September.
88. Mumpower, J.L., Livingston, S. and Lee, T.J. (1987) "Expert judgments of political riskiness," *Journal of Forecasting*, 6, 51–65.
89. Nordal, K.B. (2001) "Country risk, country risk indices and valuation of FDI: A real options approach," *Emerging Markets Review*, 2, 197–217.
90. Nowzard, B. and William, R.C. (1981) "External indebtedness of developing countries," IMF Occasional Paper 3, 3–4.
91. Oh, S. and Pedrycz, W. (2000) "Identification of fuzzy systems by means of an auto-tuning algorithm and its application to nonlinear systems," *Fuzzy Sets and Systems*, 115, 205–230.
92. Olson, D. and Shi, Y. (2005) *Introduction to Business Data Mining*, McGraw-Hill, Irwin.
93. Oral, M., Kettani, O., Cosset J.C. and Mohamed, D. (1992) "An estimation model for country risk rating," *International Journal of Forecasting*, 8, 583–593.
94. Patuwo, E., Hu, M.Y. and Hung, M.S. (1993) "Two-group classification using neural networks," *Decision Sciences*, 24, 825–845.
95. Pawlak, Z. (1982) "Rough sets," *International Journal of Information and Computer Sciences*, 11, 341-356.
96. Pawlak, Z. (1991) *Rough Sets: Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, Dordrecht.
97. Pawlak, Z. and Slowinski, R. (1994) "Rough set approach to multi-attribute decision analysis," *European Journal of Operational Research*, 72, 443–459.
98. Pentaraki, K., Doumpos, and Zopounidis, C. (1999) "An integrated approach to country risk analysis." In: *Integrating Technology & Human Decisions: Bridging into the 21st Century* (Vol. II), Proceedings of the 5th International Meeting of the Decision Sciences Institute, (Despotis, D.K. and Zopounidis, C., Eds.), New Technologies Editions, Athens, 1342–1348.
99. Press, S.J. and Wilson, S. (1978) "Choosing between logistic regression and discriminant analysis," *Journal of the American Statistical Association*, 73, 699–705.
100. Quinlan, J.R. (1993) *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, Los Altos.
101. Ramcharran, H. (1999) "The determinants of secondary market prices for developing country loans: The impact of country risks," *Global Finance Journal*, 10(2), 173–186.
102. Reinhart, C.M. (2002) "Default, Currency Crises, and Sovereign Credit Ratings," *World Bank Economic Review*, 16, 151–170.
103. Reisen, H. and Von Maltzan, J. (1999) "Boom and bust and sovereign ratings," OECD Development Centre, Technical Paper, No. 148.
104. Roy, B. (1996) *Multicriteria Methodology for Decision Aiding*, Kluwer Academic Publishers, Dordrecht.
105. Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986), "Learning internal representation by error propagation." In: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Rumelhart, D.E. and Williams, J.L., Eds.), MIT Press, Cambridge, MA.
106. Saini, K.G and Bates, P.S. (1978) "Statistical techniques for determining debt-servicing capacity for developing countries: Analytical review of the literature and further empirical results," Federal Reserve Bank of New York Research Paper No. 7818.
107. Saini, K.G. and Bates, P.S. (1984) "A survey of the quantitative approaches to country risk analysis," *Journal of Banking and Finance*, 8, 341–356.
108. Sargen, N. (1977) "Use of economic indicators and country risk appraisal," *Economic Review of the Federal Reserve Bank of San Francisco*, Fall.
109. Shapiro, A.C. (1985) "Currency risk and country risk in international banking," *Journal of Finance*, 40(3), 881–893.
110. Siskos, J. (1982) "A way to deal with fuzzy preferences in multicriteria decision problems," *European Journal of Operational Research*, 10, 314–324.
111. Siskos, Y. and Yannacopoulos, D. (1985) "UTASTAR: An ordinal regression method for building additive value functions," *Investigacao Operacional*, 5(1), 39–53.

112. Siskos, J., Lochard, J. and Lombardo, J. (1984) "A multicriteria decision-making methodology under fuzziness: Application to the evaluation of radiological protection in nuclear power plants." In: *Fuzzy Sets and Decision Analysis* (Zimmermann, H.J., Zadeh, L.A. and Gaines, B.R., Eds.), North-Holland, Amsterdam, 261–283.
113. Skowron, A. (1993) "Boolean reasoning for decision rules generation," *Lecture Notes in Artificial Intelligence*, 689, 295–305.
114. Slowinski, R. (1993) "Rough set learning of preferential attitude in multi-criteria decision making," *Lecture Notes in Artificial Intelligence*, 689, 642–651.
115. Slowinski, R. and Stefanowski, J. (1992) "RoughDAS and RoughClass software implementations of the rough sets approach." In: R. Slowinski (Ed.), *Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory*, Kluwer Academic Publishers, Dordrecht, 445–456.
116. Slowinski, R. and Stefanowski, J. (1994) "Rough classification with valued closeness relation." In: *New Approaches in Classification and Data Analysis* (Diday, E., Lechavalier, Y., Schrader, M., Bertrand, P. and Burtschy, B., Eds.), Springer-Verlag, Berlin, 482–488.
117. Slowinski, R. and Zopounidis, C. (1995) "Application of the rough set approach to evaluation of bankruptcy risk," *International Journal of Intelligent Systems in Accounting, Finance and Management*, 4, 27–41.
118. Smith, C. (1947) "Some examples of discrimination," *Annals of Eugenics*, 13, 272–282.
119. Somerville, R.A. and Taffler, R.J. (1995) "Banker judgment versus formal forecasting models: The case of country risk assessment," *Journal of Banking and Finance*, 19, 281–297.
120. Stefanowski, J. and Vanderpooten, D. (1994) "A general two-stage approach to inducing rules from examples." In: *Rough Sets, Fuzzy Sets and Knowledge Discovery* (Ziarko, W., Ed.), Springer-Verlag, London, 317–325.
121. Subramanian, V., Hung, M.S. and Hu, M.Y. (1993) "An experimental evaluation of neural networks for classification," *Computers and Operations Research*, 20(7), 769–782.
122. Subrahmaniam, K. and Chinganda, E.F. (1978) "Robustness of the linear discriminant function to nonnormality: Edgeworth series," *Journal of Statistical Planning and Inference*, 2, 79–91.
123. Taffler, R.J. and Abassi, B. (1984) "Country risk: A model for predicting debt servicing problems in developing countries," *Journal of the Royal Statistical Society*, 147(4), 541–568.
124. Tang, J.C.S. and Espinal, C.G. (1989) "A model to assess country risk," *Omega*, 17(4), 363–367.
125. Ting, W. (1988) *Multinational Risk Assessment and Management*, Greenwood Press, Westport, CT.
126. Vincke, Ph. (1976) "Une méthode interactive en programmation linèaire à plusieurs fonctions èconomiques," *RAIRO*, 10, 5–20.
127. Von Altrock, C. (1996) *Fuzzy Logic and Neurofuzzy Applications in Business and Finance*, Prentice Hall, New Jersey.
128. World Bank, (1997) *World Development Indicators 1997*, World Bank Publications, New York.
129. World Bank, (2001) *World Development Indicators 2001*, World Bank Publications, New York.
130. World Bank, (2002) *World Development Indicators 2002*, World Bank Publications, New York.
131. Yager, R.R. (1977) "Multiple objective decision-making using fuzzy sets," *International Journal of Man-Machine Studies*, 9, 375–382.
132. Zadeh, L.A. (1965) "Fuzzy sets," *Information and Control*, 8, 338–353.
133. Ziarko, W., Golan, D. and Edwards, D. (1993) "An application of DATALOGIC/R knowledge discovery tool to identify strong predictive rules in stock market data." In: Proceedings of the AAAI Workshop on Knowledge Discovery in Databases, Washington D.C., 89–101.
134. Zighed, D., Rabaseda, S. and Rakotomala, R. (1998) "FUSINTER: A method for discretisation of continuous attributes," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6(3), 307–326.

135. Zimmermann, H.J. (1978) "Fuzzy programming and linear programming with several objective functions," *Fuzzy Sets and Systems*, 1, 45–55.
136. Zopounidis, C. and Doumpos, M. (1997) "A multicriteria decision aid methodology for the assessment of country risk," *European Research on Management and Business Economics*, 3(3), 13–33.
137. Zopounidis, C. and Doumpos, M. (1998) "Developing a multicriteria decision support system for financial classification problems: The FINCLAS system," *Optimization Methods and Software*, 8, 277–304.
138. Zopounidis, C. and Doumpos, M. (1999a) "Business failure prediction using UTADIS multicriteria analysis," *Journal of the Operational Research Society*, 50(11), 1138–1148.
139. Zopounidis, C. and Doumpos, . (1999b) "A multicriteria decision aid methodology for sorting decision problems: The case of financial distress," *Computational Economics*, 14(3), 197–218.
140. Zopounidis, C. and Doumpos, M. (2000a) "PREFDIS: A multicriteria decision support system for sorting decision problems," *Computers and Operations Research*, 27(7–8), 779–797.
141. Zopounidis, C. and Doumpos, M. (2000b) "Building additive utilities for multi-group hierarchical discrimination: The MHDIS method," *Optimization Methods and Software*, 14(3), 219–240.
142. Zopounidis, C., Doumpos, M. and Zanakis, S.H. (1999) "Stock evaluation using a preference disaggregation methodology," *Decision Sciences*, 30(2), 313–336.
143. Zopounidis, C., Pentaraki, K. and Doumpos, M. (1998) "A review of country risk assessment approaches: New empirical evidence." In: *Managing in Uncertainty: Theory and Practice* (Zopounidis, C. and Pardalos, P.M. Eds.), Kluwer Academic Publishers, Dordrecht, 5–22.

Index