

Simple and multiple regression models –

A Matrix Approach

EMILIA GOGU
NICOLAE BÂRSAN-PIPU

Simple and multiple regression models –

A Matrix Approach



EDITURA UNIVERSITARĂ
București, 2017

Colecția ȘTIINȚE ECONOMICE

Redactor: Gheorghe Iovan
Tehnoredactor: Ameluța Vișan
Coperta: Monica Balaban

Editură recunoscută de Consiliul Național al Cercetării Științifice (C.N.C.S.) și inclusă de Consiliul Național de Atestare a Titlurilor, Diplomelor și Certificatelor Universitare (C.N.A.T.D.C.U.) în categoria editurilor de prestigiu recunoscut.

Descrierea CIP a Bibliotecii Naționale a României

GOGU, EMILIA

Simple and multiple regression models : a matrix approach = Modele de regresie simplă și multiplă : o abordare matriceală / Emilia Gogu, Nicolae Bârsan-Pipu. - București : Editura Universitară, 2017

Conține bibliografie
ISBN 978-606-28-0611-8

I. Bârsan-Pipu, Nicolae

311

DOI: (Digital Object Identifier): 10.5682/9786062806118

© Toate drepturile asupra acestei lucrări sunt rezervate, nicio parte din această lucrare nu poate fi copiată fără acordul Editurii Universitare

Copyright © 2017
Editura Universitară
Editor: Vasile Muscalu
B-dul. N. Bălcescu nr. 27-33, Sector 1, București
Tel.: 021 – 315.32.47 / 319.67.27
www.editurauniversitara.ro
e-mail: redactia@editurauniversitara.ro

Distribuție: tel.: 021-315.32.47 / 319.67.27 / 0744 EDITOR / 07217 CARTE
comenzi@editurauniversitara.ro
O.P. 15, C.P. 35, București
www.editurauniversitara.ro

*Modele de regresie simplă și multiplă –
O abordare matriceală*

EMILIA GOGU
NICOLAE BÂRSAN-PIPU

*Modele de regresie simplă și multiplă –
O abordare matriceală*



EDITURA UNIVERSITARĂ
București, 2017

Content

- Preface xi**
- Chapter 1: General concepts in econometrics 1**
 - 1.1 Definition and etymology of the term of econometrics..... 1
 - 1.2 Objectives and stages of econometrics 2
 - 1.3 Basic concepts..... 3
 - 1.3.1 Economic variables 3
 - 1.3.2 Types of data 3
 - 1.4 Key concepts..... 6
- Chapter 2: Econometric models..... 7**
 - 2.1 The concept of econometric model 7
 - 2.2 Types of econometric models 8
 - 2.2.1 Univariate and multivariate models 8
 - 2.2.2 Linear models and nonlinear models 9
 - 2.3 Steps of the econometric model 11
 - 2.4 Key concepts..... 14
- Chapter 3: Simple regression models 16**
 - 3.1 The components of the simple regression model 16
 - 3.1.1 The concept of "regression" 16
 - 3.1.2 Terminology for the simple regression model..... 19
 - 3.1.3 The random simple regression model 19
 - 3.2 Hypotheses of the simple linear regression model..... 20
 - 3.3 Key concepts..... 22
- Chapter 4: Estimation and inference in simple regression models 23**
 - 4.1 Matrix estimation in the simple regression model 23
 - 4.1.1 Definitions and notations 23
 - 4.1.2 Matrix estimation of simple regression model parameters..... 25
 - 4.2 Matrix inference in the simple regression model 28
 - 4.2.1 Matrix testing of hypotheses on parameters of the simple regression model 28
 - 4.2.3 *F* test 31
 - 4.2.4 Matrix analysis of residual values in the simple regression model 32
 - 4.3 Key concepts..... 33

Chapter 5: Estimation and inference in multiple regression models.....	34
5.1 Matrix estimation in the multiple regression model.....	34
5.1.1 Definitions and notations	34
5.1.2 Matrix estimation of multiple regression model parameters	36
5.2 Matrix inference in the multiple regression model	38
5.2.1 Matrix testing of hypotheses on multiple regression model parameters.....	38
5.2.2 Matrix analysis of residual values in the multiple regression model.....	47
5.3 Key concepts.....	51
Chapter 6: Nonlinear regression models.....	52
6.1 Polynomial regression models	52
6.1.1 Definitions and notations	52
6.1.2 Matrix estimation of the polynomial regression model parameters	53
6.2 Regression models with interaction	57
6.2.1 Definitions and notations	57
6.1.2 Matrix estimation of the parameters of the regression model with interaction	57
6.3 Key concepts.....	60
Chapter 7: Analysis of simple and multiple regression models in Excel.....	61
7.1 Analysis of simple regression model in Excel	61
7.1.1 The Regression option in Excel	61
7.1.2 The Linest function in Excel	64
7.2 Analysis of multiple regression model in Excel.....	65
7.2.1 The Regression option in Excel for multiple regression	65
7.2.2 The LINEST function in Excel for multiple regression	65
7.2.3 SOLVER option in Excel for multiple regression.....	66
7.3 Key concepts.....	67
Bibliography.....	68

Preface

Econometrics is a relatively new field of complex knowledge and multiple valences in the sphere of mathematics, economics, statistics and why not, of computer science. Therefore, the systemic vision of econometric research of economic phenomena requires interdisciplinarity in which analysis of interdependence of factors occupies a central place.

By combining the statistical and mathematical methods with the theoretical peculiarities of the political economy, econometrics meets both the valences of an applied economic research and the fundamental theory of general theory.

To understand the complexity and variety of phenomena, an econometric characterization is required by using arsenal of methods and analysis techniques in time, space and interdependence. In econometrics, however, this is evidenced by a specific methodology and a concrete way of interpreting or approaching the conclusions and decision-making. The information obtained through the analysis of multiple interdependencies contains mostly concrete points of view, thus facilitating the understanding of the phenomena as a whole. This is particularly important for establishing short, medium or long-term policies.

For this purpose, this paper interferes with the content and structure of a scientific and applied material, by observing the research stages, as follows:

- identify the subject, the particularities and the theoretical elements of econometrics;
- the processes, techniques and econometrics methods are iterated;
- all the stages of econometric modeling are presented and exemplified in a logical sequence;
- materializes a concrete and overall analysis of economic phenomena and processes.

The present paper can be considered as a tool for studying and applying simple and multiple regression models of economic phenomena and processes through a synthetic manner.

By content, the paper has a methodological, theoretical and applicative character. The "theory-case study" form under which the work is exposed allows, on the one hand, a better orientation of the reader in the field of economic process analysis and, on the other hand, corresponds to the logic and content of a specialized work.

The book is presented in an accessible way, in order to facilitate the understanding of the econometric mechanism in order to carry out complex studies. The information generated by

the econometric processing presented in the case studies will help readers to properly understand and evaluate the correlation between phenomena.

Through this synthesis paper, the authors aimed at presenting the econometric modeling process of data in a matrix approach, using the most appropriate methods to obtain information necessary for knowledge and decision-making.

This specialized work is addressed to researchers, statisticians and future specialists in different fields: management, quality engineering, sociology, psychology, medicine, all interested in using efficient means of analysis and knowledge of complex variables in time and space.

The authors thank those who, through suggestions and suggestions, have contributed and will contribute to the improvement of the content and quality of this specialized paper.

The authors

Cuprins

Prefață	XV
Capitolul 1: Concepte generale în econometrie	71
1.1 Definierea și etimologia termenului de econometrie	71
1.2 Obiectivele și etapele econometriei	72
1.3 Concepte de bază	73
1.3.1 Variabile economice	73
1.3.2 Tipuri de date	73
1.4 Concepte cheie	76
Capitolul 2: Modele econometrice	77
2.1 Conceptul de model econometric	77
2.2 Tipuri de modele econometrice	78
2.2.1 Modele unifactoriale și modele multifactoriale	78
2.2.2 Modele liniare și modele neliniare	79
2.3 Pașii modelului econometric	81
2.4 Concepte cheie	84
Capitolul 3: Modele de regresie simplă	85
3.1 Componentele modelului de regresie simplă	85
3.1.1 Conceptul de “regresie”	85
3.1.2 Terminologie pentru modelul de regresie simplă	88
3.1.3 Modelul aleator de regresie simplă	88
3.4 Ipotezele modelului de regresie liniară simplă	89
3.5 Concepte cheie	91
Capitolul 4: Estimare și inferență în modelele de regresie simplă	92
4.1 Estimarea matriceală în modelul de regresie simplă	92
4.1.1 Definiții și notații	92
4.1.2 Estimarea matriceală a parametrilor modelului de regresie simplă	94
4.2 Inferența matriceală în modelul de regresie simplă	97
4.2.1 Testarea matriceală a ipotezelor asupra parametrilor modelului de regresie simplă	97
4.2.3 Testul F	100
4.2.4 Analiza matriceală a valorilor reziduale în modelul de regresie simplă	101
4.3 Concepte cheie	102

Capitolul 5: Estimare și inferență în modelele de regresie multiplă.....	103
5.1 Estimarea matriceală în modelul de regresie multiplă	103
5.1.1 Definiții și notații	103
5.1.2 Estimarea matriceală a parametrilor modelului de regresie multiplă	105
5.2 Inferența matriceală în modelul de regresie multiplă.....	108
5.2.1 Testarea matriceală a ipotezelor asupra parametrilor modelului de regresie multiplă.	108
5.2.2 Analiza matriceală a valorilor reziduale în modelul de regresie multiplă	116
5.3 Concepte cheie	120
Capitolul 6: Modele de regresie neliniare	121
6.1 Modele de regresie polinomială.....	121
6.1.1 Definiții și notații	121
6.1.2 Estimarea matriceală a parametrilor modelului de regresie polinomială.....	122
6.2 Modele de regresie cu interacțiune	126
6.2.1 Definiții și notații	126
6.1.2 Estimarea matriceală a parametrilor modelului de regresie cu interacțiune	126
6.3 Concepte cheie	129
Capitolul 7: Analiza modelelor de regresie simplă și multiplă în Excel.....	130
7.1 Analiza modelului de regresie simplă în Excel	130
7.1.1 Opțiunea Regression din Excel	130
7.1.2 Funcția Linest din Excel.....	133
7.2 Analiza modelului de regresie multiplă în Excel	134
7.2.1 Opțiunea Regression din Excel pentru regresia multiplă	134
7.2.2 Funcția LINEST din Excel pentru regresia multiplă	134
7.2.3 Opțiunea SOLVER din Excel pentru regresia multiplă.....	135
7.3 Concepte cheie	136
Bibliografie.....	137

Prefață

Econometria reprezintă un domeniu relativ nou, de cunoștințe complexe și cu valențe multiple din sfera matematicii, economiei, statisticii și de ce nu, a informaticii. Prin urmare, viziunea sistemică a cercetării econometrice a fenomenelor economice necesită interdisciplinaritate în cadrul căreia analiza interdependentei factorilor ocupă un loc central.

Îmbinând metodele statistico-matematice cu particularitățile teoretice ale economiei politice, econometria întrunește atât valențele unei cercetări economice aplicative, cât și ale celei fundamentale de teorie generală.

Pentru a înțelege complexitatea și varietatea fenomenelor se impune o caracterizare econometrică prin utilizarea arsenalului de metode și tehnici de analiză în timp, spațiu și interdependență. În econometrie, toate acestea, se evidențiază printr-o metodologie proprie și modalitate concretă de interpretare sau abordare în expunerea concluziilor și luarea deciziilor. Informația obținută prin analiza multiplelor interdependențe, conține de cele mai multe ori puncte de vedere concrete, facilitând astfel înțelegerea fenomenelor în ansamblu. Acest lucru este deosebit de important pentru stabilirea politicilor pe termen scurt, mediu sau lung.

În acest scop, prezenta lucrare se interferează cu conținutul și structura unui material științific și aplicativ, prin respectarea etapelor de cercetare, astfel:

- se identifică obiectul, particularitățile și elementele teoretice ale econometriei;
- se iterează procesele, tehnicile și metodele econometrice;
- se prezintă și exemplifică, într-o succesiune logică, toate etapele modelării econometrice;
- se materializează o analiză concretă și de ansamblu a fenomenelor și proceselor economice.

Lucrarea de față, poate fi considerată un instrument de studiere și aplicare a modelelor de regresie simplă și multiplă ale fenomenelor și proceselor economice printr-o expunere sintetică.

Prin conținut, lucrarea are un caracter metodologic, teoretic și aplicativ. Forma „teorie-studiu de caz” sub care este expusă lucrarea permite, pe de o parte, o orientare mai bună a cititorului în aria analizei proceselor economice, iar, pe de altă parte corespunde logicii și conținutului unei lucrări de specialitate.

Cartea este prezentată într-un mod accesibil, în scopul ușurării înțelegerii mecanismului econometric în vederea realizării unor studii complexe. Informațiile generate prin prelucrarea econometrică, prezentate în studiile de caz, îi vor ajuta pe cititori să înțeleagă și să evalueze în mod adecvat corelație dintre fenomene.

Prin această lucrare de sinteză, autorii au avut drept scop prezentarea procesului de modelare econometrică a datelor într-o abordare matriceală, folosind cel mai adecvate metode de a obține informații necesare cunoașterii și luării deciziilor.

Această lucrare de specialitate se adresează cercetătorilor, statisticienilor și viitorilor specialiști din domenii diferite: management, ingineria calității, sociologie, psihologie, medicină, tuturor celor interesați în utilizarea mijloacelor eficiente de analiză și cunoaștere a fenomenelor complexe variabile în timp și spațiu.

Autorii mulțumesc celor care, prin propuneri și sugestii, au contribuit și vor contribui la îmbunătățirea conținutului și calității prezentei lucrări de specialitate.

Autorii

Chapter 1

General concepts in econometrics

1.1 Definition and etymology of the term of econometrics

Econometrics, whose name originates in the Greek word combination, *eikonomia* (economy) with *metron* (measure) has a multiseccular evolution.

Econometrics literally means economic measurements. It is a combination of economics, mathematics and statistics.

To begin with, we will present and analyze the main definitions of econometrics given in dictionaries, but we would like to point out that in the literature there are dozens of definitions of econometrics in that it is a somewhat new and developing science.

Econometrics (f. s.): The set of mathematical and statistical methods used as a tool for studying the quantitative correlations of economic phenomena and processes. From fr. économétrie. Source: Explanatory Dictionary of Romanian Language (DEX).

Econometrics (f.): Branch of the economy dealing with the study of methods of mathematical and statistical analysis of economic phenomena and processes.

Source: The New Explanatory Dictionary of Romanian Language (NODEX).

Econometrics: Statistical and mathematical analysis of economic relations. Econometrics creates equations to describe phenomena such as the relationship between price changes and demand.

The term *econometrics* was introduced by Ragnar Frisch, a Norwegian economist and statistician, the first Nobel laureate for the economy. In a work in 1926, Frisch wrote: "As an intermediary between mathematics, statistics, and economics, we will find a new discipline that can be called econometrics. Econometrics aims to obey the abstract laws of the theoretical political economy by experimental and numerical verification, and thus to transform the pure economy, as far as possible, into science in the strict sense of the word. "

Econometrics is an economic science that represents a coherent set of methods and instruments of statistical analysis - mathematics of phenomena and economic processes in which economic theory itself is formalized in mathematical terms. By combining the statistical and mathematical methods with the theoretical particularities of the political economy, econometrics brings together both the valences of an applied and the fundamental theory of general theory.

1.2 Objectives and stages of econometrics

In general, modern approaches to the economy can be classified into the following categories: *macroeconomics*, *microeconomics*, *financial economics* and *econometrics*. In all these economic categories, a key role is the quantitative analysis, a role that econometrics performs.

In this context, the main objectives of econometrics can be synthesized as follows:

- a. ***Estimation of economic relations***. Analysis of how economic theory can explain the available data about a particular economic process.
- b. ***Confronting the economic theory with reality and testing hypotheses about aspects specific to the studied phenomenon***. Testing the validity of economic theory and economic model assumptions.
- c. ***Forecasting economic variables***. Forecast on the evolution of the analyzed economic process.

Broadly speaking, to achieve these objectives, econometric analysis studies should go through the following stages:

Stage 1: *Data collection and synthesis of observed events and trends regarding the economic processes under review*. This stage concerns the focus on certain economic processes, for which we observe and collect and record data and facts. These data may be *income and consumption trends, financial market trends and investment volume, inflation rate developments, export / import growth / decrease and unemployment rate dynamics*, etc.

Stage 2: *Development of the econometric model and formulation of the theoretical model in mathematical format*. Based on the collected data and empirical observations on the economic process, at this stage, it is necessary to construct an econometric model, explaining the economic phenomenon analyzed. This implies the development of a statistical-mathematical model, to be transformed into an econometric model. In general, the economic model only provides us with a quantitative assessment without having a specific functional relationship. This relationship of quantitative nature is now determined by determining the model variables and the functional relationships between these variables.

Stage 3: *Checking the econometric model*. For the determined econometric model, it is necessary to estimate the unknown parameters of the model based on the recorded data. Next, check whether the econometric model is appropriate, or whether it is consistent with the model formulated assumptions.

Stage 4: *Applying the econometric model*. After the verification of the econometric model, it can be applied to validate the economic model of the process under review, to make forecasts and scenarios on future developments and to substantiate economic and financial management decisions.

1.3 Basic concepts

1.3.1 Economic variables

The *economic variables* determine the structure of the econometric model: exogenous variables or endogenous variables.

The *random variable* (ε) synthesizes the totality of the variables (apart from the factorial ones) that influence the endogenous variable but are not specified within the model (random factors).

The *time variable* (t) is introduced into some econometric models as a fictitious variable.

The variables of the econometric model are as follows:

- 1) The *input variables*, X_1, X_2, \dots, X_k , also called *independent variables* or *explanatory variables* or *factorial variables* or *exogenous variables*, are random / determined variables outside the system, about which the econometric model has nothing to say;
- 2) The *output variable*, Y , also called the *dependent variable* or the *explained variable* or the *resultative variable* or the *endogenous variable*, which is also a random variable / determined within the system;
- 3) The *perturbation variables*, denoted by u or ε , also called *residual variables* or *error variables*, synthesize the totality of the variables (apart from the factorial ones) that influence the endogenous variable but are not specified within the model (random factors);
- 4) The *time variable* (t) is introduced in some econometric models as a fictitious variable for two reasons:
 - Allows to identify regularities in the evolution of phenomena;
 - Represents the artificial measure of those economic variables that act on the resulting variable but which, being of a qualitative nature, can not be quantified or appear explicitly in the model.

1.3.2 Types of data

From a chronological point of view, the observation of economic phenomena and processes can be done statically or dynamically, evolutionarily. These two ways of observing the units of a population lead to the grouping of statistical data into three categories.

The types of data used in econometric models classified as follows:

- *Cross-sectional data*;
- *Time-series data*;
- *Longitudinal* or *panel data*.

- a) **Section / profile type cross-sectional data, at statistical unit level**, are data recorded for one or more econometric variables, for a single entity / characteristic, obtained at a given time, respectively in a single time period). There are transversal information sections, called "information cuts" in relation to the time axis.

Ex: Level of exports to EU countries in 2016.

The entities may be individuals, consumers, manufacturers, firms, government bodies, states, etc. For cross-sectional data, we'll track them through aliasing sampling methods. Cross-sectional data allow us to establish relationships between the econometric model variables for the entity at a given point in time. The transversal data recording model is presented in the table below:

<i>Observation</i>	<i>Dependent variable</i>	<i>Independent variables</i>				
i	Y_i	X_{1i}	X_{2i}	...	X_{ki}	
1	Y_1	X_{11}	X_{21}	...	X_{k1}	
2	Y_2	X_{12}	X_{22}	...	X_{k2}	
...	
i	Y_i	X_{1i}	X_{2i}	...	X_{ki}	
...	
n	Y_n	X_{1n}	X_{2n}	...	X_{kn}	

- b) The **time series data** (chronological series) are the data recorded for one or more econometric variables for a single entity (individual, firm, country) collected over several discrete time periods (years, quarters, months, weeks, days etc.). There are longitudinal information sections, called "information sections" in relation to the time axis.

Ex. The level of exports of goods and services in the period 2000-2015, for Romania.

The main issue of this type of data is that in most cases they are not independent over time, most of which are correlated over time. Time series data allow us to track the evolution of the econometric model variables and to produce forecasts for that entity at certain future times. The time series data recording model is presented in the following table:

<i>Observation</i>	<i>Period of time</i>	<i>Dependent variable</i>	<i>Independent variables</i>			
i	t	Y_i	X_{1i}	X_{2i}	...	X_{ki}
1	t_1	Y_1	X_{11}	X_{21}	...	X_{k1}
2	t_2	Y_2	X_{12}	X_{22}	...	X_{k2}
...
i	t_i	Y_i	X_{1i}	X_{2i}	...	X_{ki}
...
n	t_n	Y_n	X_{1n}	X_{2n}	...	X_{kn}