Ministry of education of the Republic of Belarus

Educational institution
"Pavel Sukhoi Gomel State Technical University"

Department of "Marketing and industry economics"

O. G. Vinnik

ECONOMETRICS

TEXTBOOK
for specialty 6-05-0311-02 "Economics and management"
for full-time students

Gomel 2024

UDS 519.8(075.8)
BBK 65в631я73
V69

Recommended for publication by the Sukhoi GSTU council
(*protocol No 8 dates 01.03.2024*)

Reviewer: dean of the faculty of Humanities and Economics of Educational establishment “Pavel Sukhoi Gomel State Technical University”, Ph.D., associate professor *M. N. Andrianchikova*

Vinnic, O. G.
V69 Econometrics : textbook for specialty 6-05-0311-02 “Economics and management” for full-time students / O. G. Vinnik. – Gomel : GSTU, 2024. – 48 p. – Системные требова-ния: PC не ниже Intel Celeron 300 МГц ; 32 Mb RAM ; свободное место на HDD 16 Mb ; Windows 98 и выше ; Adobe Acrobat Reader. – Режим доступа: https://elib.gstu.by. – Загл. с титул. экрана.

The basic categories and concepts of international financeare outlined in the textbook, international financial flows, participants, main types of markets ans and transactions as well. The textbook gives the understanding of concepts of international finance.
The textbook can be recommended for students of economic specialties.

UDS 519.8(075.8)
BBK 65в631я73

CONTENT

INTRODUCTION

The relevance of econometrics as an academic discipline assumes high significance because of the interdisciplinary nature of econometrics. In the ongoing scenario of globalization and economic deregulation, there is the need to give added thrust to the academic discipline of econometrics in higher education. Accordingly, the analytical ability of the students can be sharpened and their ability to look into the socio-economic problems with a mathematical approach can be improved, and enabling them to derive scientific inferences and solutions to such problems. The utmost significance of hands-on practical training on the use of computer-based econometric packages. Learning of the econometrics have much practical utility for the students in their future career, whether in academics, industry, or in practice.

For students studying Econometrics in a university it is relevance, whether they decide to start research career after receiving PhD degree in Economics or wish to start career outside science. For future researchers Econometrics provides a strong analytical background and prepare them for more advanced research activity, for future finance specialists a module in Econometrics provide specific and common skills useful in finance area.

TOPIC 1. INTRODUTION TO DISCIPLINE. ECONOMETRIC MODELS AND METHODS IN RESEARCH.

Plan:
1. Subject and content of the course.
2. Theoretical foundations of economic and mathematical modeling.

1. Subject and content of the course

The *purpose* of teaching the discipline is to teach students the basics of economic and mathematical modeling, as well as methods of development and implementation of models in the process of production organization, economic analysis, planning and forecasting.

The main objectives of the discipline are studying:

- introduction to the theoretical foundations of econometrics, economic and mathematical modeling and operations research;

- foundations of regression analysis and its using in economics of organization, region or country studying;

- methods of setting tasks;

- mastering the skills of finding solutions to economic and mathematical optimization problems;

- game theory methods.

2. Theoretical foundations of economic and mathematical modeling

In the today's economics, mathematical modeling is a very important research and forecasting tool for practical and theoretical problems. In modern economy are widely used various mathematical methods for solving practical problems, and theoretical modeling of socio - economic processes.

Mathematical modeling is ideal scientific symbol-based formal modeling, which produces subject descriptions in the language of math, while the models are researched by means of mathematical formulas.

Mathematical economics is an approach to economic analysis where mathematical symbols and theorems are used.

The Mathematical Programming Problem

The basic problem of economics, economizing, is that of allocating scarce resources among competing ends. Because of the scarcity of resources, choices must be made, and rational choices are those attaining certain objectives within the limitation of resource scarcity.

The **mathematical problem in the language of economics** is the following: how to choose the instruments within the opportunity set so as to maximize or minimize the objective function.

In the course of study, economic and mathematical modeling is, for the most part, identical to the study of operations.

In operations research, problems are broken down into basic components and then solved in defined steps by mathematical analysis.

Operations Research assists decision-makers in almost any management function.

1. Manufacturing.
2. Revenue Management.
3. Supply Chain Management.
4. Retailing.
5. Financial Services.
6. Marketing Management.
7. Human Resource Management.
8. General Management
9. Production systems.

Operation research models are designed to optimize a specific objective criterion subject to a set of constraints, the quality of the resulting solution depends on the degree of completeness of the model in representing the real system.

The most prominent Operation research technique is linear programming. Other techniques include integer programming, dynamic programming, network programming, and nonlinear programming. These are only a few among many available Operation research tools.

Some mathematical models may be so complex that it becomes impossible to solve them by any of the available optimization algorithms. In such cases, it may be necessary using heuristics rules that move the solution point advantageously toward the optimum.

TOPIC 2. THE CONTENT OF ECONOMIC AND MATHEMATICAL MODELS AND METHODS OF THEIR CONSTRUCTION.

Plan:

1. The concept of "model" and "simulation". Classification of mathematical models.
2. The main stages of economic and mathematical simulation. Principles of construction of economic and mathematical models.

3. Methodology of Operation Research, Operation Research tools.
4. The main types of distribution problems: the actual distribution problem, the problem of blending, the transport problem, the appointment.

<u>1. The concept of "model" and "modeling". Classification of models</u>

The essence of the operations research activity lies in the construction and use of models.

When a problem or process under investigation is simplified and represented with its typical features or characteristics, it is called as a **model**.

A **model in the sense used in operation research** is defined as a representation of an actual object or situation. It shows the relationships and reaction in terms of cause and effect.

Models describe our beliefs about how the world functions. In mathematical modelling, we translate those beliefs into the language of mathematics.

Simulation – is the imitative representation of the functioning of one system or process by means of the functioning of another.

Why simulate?

- It may be too difficult, hazardous, or expensive to observe a real, operational system;
- Parts of the system may not be observable.

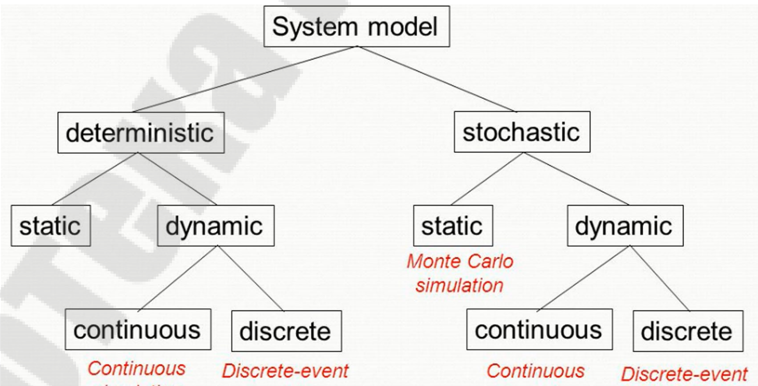


Fig. 1. Types of Simulation Models

<u>Deterministic models</u>

•Assume that the outcome is precisely determined by the model inputs and relationships

•Ignore all random variation

•A given input always produces the same output

Stochastic models

•Incorporate inherent randomness

•Use a range of values for the model variables in form of probability distributions

•The same input produces an ensemble of outputs

Static models

- static model is the model of the system not during runtime
- the doesn't of differential equations
- static models are at equilibrium of in a steady state
- is more structural
- includes class diagram and object diagrams and help in depicting static constituents of the system
- It cannot be changed in real time

Dynamic models

- dynamic model refers to runtime model of the system
- the use of differential equations in dynamic model
- keep changing with reference to time
- is a representation of the behavior of the static components of the system
- consists of sequence of operations, state changes, activities, interactions and memory
- can change with time as it shows

2. The main stages of economic and mathematical modeling. Principles of construction of economic and mathematical models

Mathematical modeling involves two equally important activities:

• building a mathematical structure, a model, based on hypotheses about relations among the quantities that describe the real world situation, and then deriving new relations;

• evaluating the model, comparing the new relations with the real world and making predictions from the model.

Important principles of model building

• Models should not contradict fundamental laws of nature

• Testing (validation) of models against basic laws of nature

• Proportionality should be taken into account

• Scaling can be exploited to reduce the complexity

• To use analogies (e.g. chemical reactions and competition models)

• Universality: different objects are described by the same model (e.g. vibrations of the body of a car and the passage of signals through electrical filters)

• Hierarchical principle: each model may incorporate submodels – a step-like refinement

• Modularity and reusability

• From problem to method, not vice versa

3. Methodology of Operation Research:

1. Formulating the Problem: (or Definition of the problem and Construction of the model);
2. Deriving Solution (or Solution of the model);
3. Testing (Validation) the Model and Solution;
4. Establishing Controls over the Solution;
5. Implementation and Implementing the Solution.

The most common operation research tool is Linear Optimization, or Linear Programming (LP).

The "Programming" in Linear Programming is synonym for "optimization". It has nothing to do with computer programming. Most real-world problems are "close enough" to linear problems.

Constructing the Model

A **mathematical model** is a set of equations in which the system or problem is described. The equations represent objective function and constraints.

Objective function is a mathematical expressions of objectives (cost or profit of the operation), while **constraints** are mathematical expressions of the limitations on the fulfillment of the objectives.

The general Linear Programming Problem calls for optimizing (maximizing/minimizing) a linear function for variables called the 'objective function' subject to a set of linear equations and/or inequalities called the 'constraints or restrictions.'

4. The main types of distribution problems

Basic Requirements and their Relationships:

1. Decision Variables and their Relationships.
2. Objective Function.
3. Constraints
4. Alternative Courses of Action.
5. Non-Negativity Restrictions.

6. Linearity and Divisibility.
7. Deterministic.

Classical LP problem. *Example*: A company produces both interior and exterior paints from two raw materials, M_1 and M_2. The following table provides the basic data of the problem:

Table 1

	Tons of raw material per ton of		Maximum daily availability (tons)
	exterior paint	interior paint	
Raw material, M_1	6	4	24
Raw material, M_2	1	2	6
Profit per ton ($1000)	5	4	

The daily demand for interior paint cannot exceed that for exterior paint by more than 1 ton. Also, the maximum daily demand for interior paint is 2 tons. A company wants to determine the optimum (best) product mix of interior and exterior paints that maximizes the total daily profit.

Formulate the problem as LPP.

Solution:

Classical LP problem an algebraic Model

$$maximize\ z = 5x_1 + 4x_2$$

subject to restriction

$$6x_1 + 4x_2 \le 24$$
$$2x_1 + 2x_2 \le 6$$
$$x_1 - x_2 \le 1$$
$$x_2 \le 2$$
$$x_1, x_2 \ge 0$$

A company problem — In Σ Sigma notation

$$maximize\ z = \sum_{j=1}^{n} c_j \times x_j$$

subject to

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i,\ i=1,\ 2,\ \dots,\ m$$

$$x_j \geq 0,\ j=1,\ 2,\ \dots,\ n$$

where i - raw material index;

j – paint type index

m – number of raw materials (m=2);

n – number of raw materials (n=2);

c_j – profit per ton ($1000) of j-type paint;

a_{ij} – tons of i-type raw material per ton of of j-type paint;

b_i – maximum daily availability (tons) of i-type raw material.

Generalized Form of a Classical LP Problem

$$maximize\ z = c_1x_1 + c_2x_2$$

subject to

$$a_{11}x_1 + a_{12}x_2 \leq b_1$$
$$a_{21}x_1 + a_{22}x_2 \leq b_2$$
$$x_1,\ x_2 \geq 0.$$

Diet Problem. *Example.* A farmer uses at least 800 pounds (lb) of special feed daily. The special feed is a mixture of corn and soybean meal with the following compositions:

Table 2

Feedstuff	The components of feedstuff, lb (фунт) per lb of feedstuff		Cost ($/lb)
	protein	fiber	
Corn	.09	.02	.3
Soybean meal	.60	.06	.9

The dietary requirements of the special feed are at least 30% protein and at most 5% fiber.

The goal is to determine the daily minimum-cost feed mix.

Solution:

Diet Problem an algebraic Model

$$minimize\ z = .3x_1 + .9x_2$$

subject to

$$x_1 + x_2 \geq 800$$
$$(.09x_1 + .6x_2) \leq .3\,(x_1 + x_2) \quad \text{or} \quad .21x_1 - .3x_2 \leq 0$$
$$.02x_1 + .06x_2 \leq .05\,(x_1 + x_2) \quad \text{or} \quad .03x_1 - .01x_2 \geq 0$$
$$x_1, x_2 \geq 0$$

Diet problem — In Σ Sigma notation

$$minimize\ z = \sum_{i=1}^{m} c_i \times x_i$$

subject to

$$\sum_{i=1}^{m} x_i \geq 800 \text{ (the daily amount of the mix)}$$
$$\sum_{i=1}^{m} a_{ij} x_i \leq b_j \sum_{i=1}^{m} x_i, \quad j=1, 2, \dots, n \text{ (the dietary requirements)}$$
$$x_i \geq 0,\ i=1, 2, \dots, m$$

where i - feedstuff index;
j – component index;
m – number of feedstuffs (m=2);
n – number of components (n=2);
c_i – cost ($/lb) of i-th feedstuff;
a_{ij} – the j-th component content per lb of i-th feedstuff;
b_j – the part of j-th component in the daily amount of the mix.

Generalized Form of a Diet Problem

$$minimize\ z = c_1 x_1 + c_2 x_2$$

subject to

$$x_1 + x_2 \geq 800$$
$$a_{11}x_1 + a_{12}x_2 \leq b_1(x_1 + x_2)$$
$$a_{21}x_1 + a_{22}x_2 \leq b_2(x_1 + x_2)$$
$$x_1, x_2 \geq 0.$$

Transportation problem. *Example.* MG Auto has three plants in Los Angeles, Detroit, and New Orleans and two major distribution centers in

Denver and Miami. The quarterly capacities of the three plants are 1000, 1500, and 1200 cars, and the demands at the two distribution centers for the same period are 2300 and 1400 cars. The transportation costs ($) per car on the different routes are given in Table 3.

Table 3

Transportation costs ($) per car

	Denver	Miami
Los Angeles	80	215
Detroit	100	108
New Orlean	102	68

Task: to determine the number of cars to be shipped from each plant to each distribution center so as lowest transportation cost.

Solution:

Add the missing data(plant capabilities and distribution center needs) to the source data table (with transportation costs).

a_i / b_j	2300	1400
1000	80	215
1500	100	108
1200	102	68

First, we need to determine whether this transport task is open or closed.
the capabilities of all suppliers are 1000 + 1500 + 1200 = 3700
the needs of all consumers are 2300 + 1400 = 3700

Opportunities are equal to needs, hence a closed-type task. If this was an open-type task, we would have to make it closed.

Restrictions in a transport task are always constructed with the "=" sign.

__Transportation Problem an algebraic Model__

$$minimize\ z = 80x_{11} + 215x_{12} + 100x_{21} + 108x_{22} + 102x_{31} + 68x_{32}$$

subject to

$$x_{11} + x_{12} = 1000$$
$$x_{21} + x_{22} = 1500$$

$$x_{31} + x_{32} = 1200$$
$$x_{11} + x_{21} + x_{31} = 2300$$
$$x_{12} + x_{22} + x_{32} = 1400$$
$$x_{ij} \geq 0$$

Transportation problem — In Σ Sigma notation

$$minimize\ z = \sum_{i=1}^{m} c_{ij} \times x_{ij}$$

subject to

$$\sum_{i=1}^{m} x_{ij} = b_j \text{ (needs of consumers)}$$

$$\sum_{j=1}^{n} x_{ij} = a_i \text{ (capacity of suppliers)}$$

$$x_{ij} \geq 0,$$

$$j=1, 2, \ldots, n$$
$$i=1, 2, \ldots, m$$

where i - supplier index; j – consumer index;

m – number of suppliers (m=3);

n – number of consumers (n=3);

x_{ij} – the number of cars to be shipped from *i-th* supplier (each plant) to *j-th* consumer (each distribution center);

c_{ij} – the transportation costs (\$) per car from *i-th* supplier (each plant) to *j-th* consumer (each distribution center);

a_i – the capacity of *i-th* supplier.

b_j – the needs of *j-th* consumers.

Allocation (Assignment) problem. These models are used to solve the problems arising when:

(a) There are number of activities which are to be performed and there are number of alternative ways of doing them,

(b) The resources or facilities are limited, which do not allow each activity to be performed in best possible way. Thus these models help to combine activities and available resources so as to optimise and get a solution to obtain an overall effectiveness.

Example. A manager has three workers, A, B and C, who are to be assigned to three jobs, 1, 2 and 3. The alternatives and the estimated job completion times in days are given in the table below.

	1	**2**	**3**
A	17	10	12
B	9	8	10
C	14	4	7

The manager wishes to minimise the total number of days required to complete all three jobs. How should the allocation of workers to jobs be made?

Solution:

The assignment problem is a special case of the transportation problem in which:

(i) supply = 1

(ii) demand = 1

(iii) amount shipped over each arc is either 0 or 1.

Suppose x_{jj} is a variable which is defined as

$$x_{ij} = \begin{cases} 1, & \text{if worker } i \text{ does job } j \\ 2, & \text{otherwise} \end{cases}$$

Assignment Problem an algebraic Model

$$minimize\ z = 17x_{11} + 10x_{12} + 12x_{13} + 9x_{21} + 8x_{22} + 10x_{23} + 14x_{31} + 4x_{32} + 7x_{33}$$

subject to

for A: $x_{11} + x_{12} + x_{13} = 1$

for B: $x_{21} + x_{22} + x_{23} = 1$

for C: $x_{31} + x_{32} + x_{33} = 1$

for 1: $x_{11} + x_{21} + x_{31} = 1$

for 2: $x_{12} + x_{22} + x_{32} = 1$

for 3: $x_{13} + x_{23} + x_{33} = 1$

<u>Assignment problem — In Σ Sigma notation</u>

$$minimize\ z = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij} \times x_{ij}$$

subject to

$$\sum_{i=1}^{m} x_{ij} = 1$$

$$\sum_{j=1}^{n} x_{ij} = 1$$

x_{ij} is either 0 or 1.

$j=1, 2, \ldots, n$

$i=1, 2, \ldots, m$

where i - worker index;

j – job (machine) index;

m – number of worker (m=3);

n – number of job (machine) (n=3);

c_{ij} – the number of days the *i-th* worker needs to complete the *j-th* jobs;

x_{ij} – the *i-th* worker has been assigned to *j-th* job (machine);

a_i – the capacity of *i-th* supplier.

b_j – the needs of *j-th* consumers.

TOPIC 3. METHODS FOR SOLVING OPTIMIZATION ECONOMIC AND MATHEMATICAL PROBLEMS

<u>Plan:</u>

1. Graphical method for solving linear programming problem.
2. Simplex method.
3. Methods for solving of transportation type problems.

<u>1. Graphical method for solving linear programming problem</u>

The graphical solution includes follows steps:

Step I. Defining the problem.

Step II. Since the two decision variable x_1 and x_2 are non-negative, consider only the first quadrant of *xy*-coordinate plane.

Step III. Plot the constraints Graphically. Each inequality in the constraint equation has to be treated as an equation. An arbitrary value is assigned to one variable & the value of the other variable is obtained by solving the equation. In the similar manner, a different arbitrary value is again assigned to the variable & the corresponding value of other variable is easily obtained. These 2 sets of values are now plotted on a graph and connected by a straight line. The same procedure has to be repeated for all the constraints. Hence, the total straight lines would be equal to the total no of equations, each straight line representing one constraint equation.

Step IV. Locate the solution space. Solution space or the feasible region is the graphical area which satisfies all the constraints at the same time. Such a solution point (x, y) always occurs **at the comer**. points of the feasible Region the feasible region is determined as follows: (a) For "greater than" & "greater than or equal to" constraints (i.e.;), the feasible region or the solution space is the area that lies above the constraint lines. (b) For "Less Then" &"Less than or equal to" constraint (ie;). The feasible region or the solution space is the area that lies below the constraint lines.

Step V. Identify the feasible solution region. The feasible solution region represents the area on the graph that is valid for all constraints. Choosing any point in this area will result in a valid solution.

Step VI. Plot two objective function lines to determine the direction of improvement. Improvement is in the direction of greater value when the objective is to maximize the objective function, and is in the direction of lesser value when the objective is to minimize the objective function. The objective function lines do not have to include any of the feasible region to determine the desirable direction to move.

Step VII. Find the most attractive corner. Optimal solutions always occur at corners. The most attractive corner is the last point in the feasible solution region touched by a line that is parallel to the two objective function lines drawn in step 5 above. When more than one corner corresponds to an optimal solution, each corner and all points along the line connecting the corners correspond to optimal solutions. We'll use an example to illustrate optimal solutions later.

Step VIII. Determine the optimal solution by algebraically calculating coordinates of the most attractive corner.

Step IX. Determine the value of the objective function for the optimal solution.

Solution of a Maximization Model

$$\textit{maximize } z = 5x_1 + 4x_2$$

subject to

$$6x_1 + 4x_2 \leq 24 \ (1)$$
$$2x_1 + 2x_2 \leq 6 \ (2)$$
$$x_1 - x_2 \leq 1 \ (3)$$
$$x_2 \leq 2 \ (4)$$
$$x_1, x_2 \geq 0 \ (5)$$

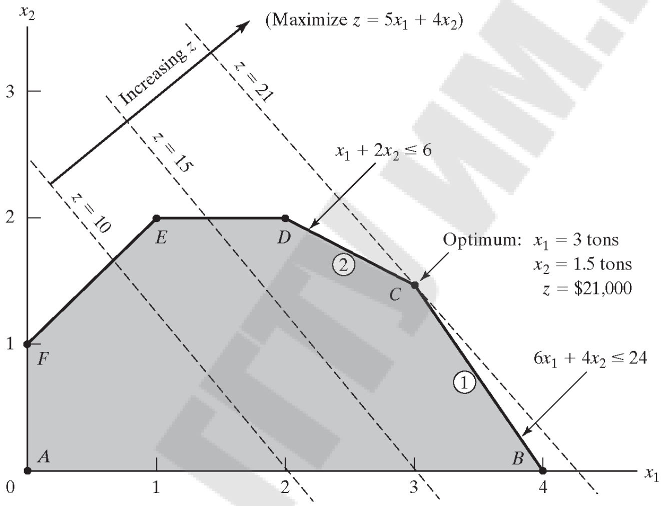


Fig. 2. Optimum solution of the model

Solution of a Minimization Model

$$\textit{minimize } z = .3x1 + .9x2$$

subject to

$$x_1 + x_2 \geq 800$$
$$.21x_1 - .30x_2 \leq 0$$
$$.03x_1 - .01x_2 \geq 0$$
$$x_1, x_2 \geq 0$$

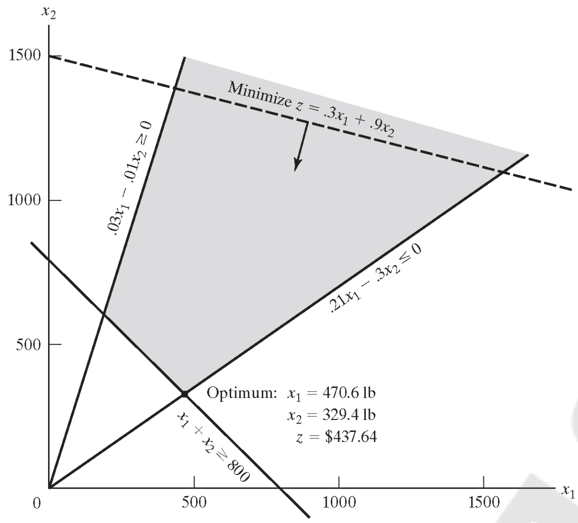


Fig. 3. Optimum solution of the model

2. Simplex method

The simplex method solution includes follows steps:

1. Convert the inequalities into equalities by adding slack variables, surplus variables or artificial variables, as the case may be.

Type of inequality	**Type of variable**
$\geq$	- surplus + artificial
=	+ artificial
$\leq$	+ slack

2. Tabulate the data into the first iteration of Simplex Method.

(a) C_j is the coefficient of unknown quantities in the objective function.

(b) $Z_j - C_j = \sum C_{bi} \cdot a_{ij} - C_j$

Table 4

General view of Simplex Method iteration

Base	C_{bi}	x_1	x_2	…	x_n	b_i	Minimum ratio b_i / a_{ij} $(a_{ij}>0)$
		C_1	C_2	**…**	C_n		
S_1	C_{b1}	a_{11}	a_{12}	…	a_{1n}	b_1	
S_2	C_{b2}	a_{21}	a_{22}	…	a_{2n}	b_2	
…	…	…	…	…	…	…	
S_m	C_{bm}	a_{m1}	a_{m2}	…	a_{mn}	b_m	
Zj - Cj	**-**	Z_1-C_1	Z_2-C_2	**…**	Z_n-C_n	Z_0	

3. Checking of the simplex optimality condition.

4. If we haven't reached the problem's optimal solution:

(a) Identify the Key or **Pivotal column** with the minimum element of Z_j - C_j denoted as '**KC**' throughout to the problems in the chapter.

(b) Find the 'Minimum Ratio' i.e., b_i/a_{ij}.

(c) Identify the **key row** with the minimum element in a minimum ratio column. Key row is denoted as '**KR**'.

(e) Identify the **key element** at the intersecting point of key column and key row, which is put into a box throughout to the problems in the chapter.

4. Reinstate the entries to the next iteration of the simplex method.

(a) The pivotal or key row is to be adjusted by making the key element as '1' and dividing the other elements in the row by the same number.

New Pivot Row Element = Actually Pivot Row Element / Pivot. (3.1)

(b) The key column must be adjusted such that the other elements other than key elements should be made zero.

(c) The same multiple should be used to other elements in the row to adjust the rest of the elements. But, the adjusted key row elements should be used for deducting out of the earlier iteration row.

The left elements of rows will be reached so:

*New Row Element = Actually Pivot Row Element - (Pivot Column Element from actually Row * New Row Element).* (3.2)

(d) The same iteration is continued until the values of $Z_j – C_j$ become either '0' or positive.

5. Find the '*Z*' value given by CB, XB.

TOPIC 4. INTRODUCTION TO ECONOMETRICS

Plan :

1. Theoretical foundations of econometrics.
2. Types and features of Econometric Models.
3. Types of Econometrics data.

1. 1. Theoretical foundations of econometrics

Econometrics is the application of mathematics, statistical methods, and, more recently, computer science, to economic data and is described as the branch of economics that aims to give empirical content to economic relations.

Econometrics is the branch of economics concerned with the empirical estimation of economic relationships, models, together with data, represent the basic ingredients of any econometric study.

Econometrics is the field of economics that concerns the application of mathematical statistics and the tools of statistical inference to the empirical measurement of relationships postulated by economic theory.

Econometrics is the quantitative application of statistical inferences, economic theory and mathematical models using data to develop theories or test existing hypotheses in economics and to forecast future trends from the huge amount of data acquired over time.

Why study Econometrics?

- Rare in economics (and many other areas without labs!) to have experimental data;
- Need to use nonexperimental, or observational, data to make inferences;
- Important to be able to apply economic theory to real world data.
- An empirical analysis uses data to test a theory or to estimate a relationship;
- A formal economic model can be tested;
- Theory may be ambiguous as to the effect of some policy change – can use econometrics to evaluate the program.

Econometrics object is an economics of spaces of different levels of complexity, from a single company or firm in the economy sectors, regions, states and the world as a whole.

Econometrics subject are methods of construction and study of mathematical and statistical models of the economy, carrying out quantitative studies of economic phenomena, explanation and forecasting of economic processes.

Purpose of econometrics research is analysis of real economic systems and processes which are taking place, using econometric methods and models and their application in making science-based management decisions.

The main task of econometrics - estimate model parameters on the base of input information features to verify the compliance model of the phenomenon and prediction of economic processes.

Features of Econometric models are following:

- **Statistical dependency** means relationship, when you change from one random variable changes the probability distribution law of another;
- **Correlation dependency** means dependence, which manifests itself in the fact that you change from one value changes the average value of the other
- **Regression equation (models)** is relations between the dependent and independent variables that are described by the relationships (formulas 4.1, 4.2).

$$y = f(x) + \varepsilon, \tag{4.1}$$

$$y = F(x_1, x_2, \dots, x_m) + \varepsilon. \tag{4.2}$$

The two main types of regression are as follows:

– *Pair* – relationship of two random variables, and

– *Multiple* - relationship of *y* t and several random variables,

2. <u>Types of econometric models</u>

The main tool of econometrics are the simple and multiple linear regression model.

A **model** is a simplified representation of a real world process. It should be representative in the sense that it should contain the salient features of the phenomena under study.

Econometrics uses statistical methods after adapting them to the problems of economic life.

Econometric model - a function or system of functions that describes the correlation and regression relationship between economic indicators and on the base of the causal links between one or more of these parameters we can determine some variables as dependent variables, and others - as independent ones.

<u>The reasons for the presence of random factors in regression models</u>

1. Putting into model not all explanatory variables.
2. The wrong choice of functional form model.
3. Aggregation of variables.
4. Errors of measurement.
5. The limited statistics.
6. The unpredictability of the human factor.

Econometric models can be classified in a ***categories***:

1. ***A first class*** of models describes relationships between present and past. This type of models is called time series models. They are mainly used for built to get forecasts for future values.

2. ***A second class*** of models considers relationships between economic quantities over a certain time period. These relationships give us information on how economic quantities fluctuate over time in relation to other quantities.

3. ***A third class*** of models describes relationships between different variables measured at a given point in time for different units (for example households or firms). Most of the time, this type of relationships is meant to explain why these units are different or behave differently.

3. The types of Econometrics data

There are two types of econometrics data: *quantitative* and *qualitative*.

Quantitative data is data that can be expressed as a number or can be quantified. In other words, quantitative data can be measured by numerical variables.

Quantitative data are easily amenable to statistical manipulation and can be represented with a wide variety of statistical types of graphs and chards such as line, graph, bar graph, scatter plot, box and whisker plot and etc.

There are two general types of quantitative data:

Discrete data – a count that involves integers. Only a limited number of values is possible. The discrete values cannot be subdivided into parts. For example, the number of children in a school is discrete data.

Continuous data – information that could be meaningfully divided into finer levels. It can be measured on a scale or continuum and can have almost any numeric value. For example, you can measure your height at very precise scales — meters, centimeters, millimeters and etc.

Qualitative data is information that can't be expressed as a number and can't be measured.

Qualitative data consist of words, pictures, observations, and symbols, not numbers. It is about qualities.

Qualitative data is also called categorical data. The reason is that the information can be sorted by category, not by number. Qualitative data is analyzed to look for common themes.

For the purposes of constructing an econometric model, **several types of empirical data** are distinguished

– *Cross-Sectional Data*

Cross-sectional data refers to information collected from different individuals, firms, or regions at a specific point in time. It provides a snapshot of a population at a given moment.

– *Time Series Data*

Time series data is one of the most commonly used types of data in econometrics. It refers to data collected over a specific period of time at regular intervals.

– *Panel Data*

Panel data combines elements of both time series and cross-sectional data. It involves collecting information from multiple individuals or entities over a period of time.

TOPIC 5. AN INTRODUCTION TO SIMPLE REGRESSION

Plan:

1. The types of econometric variables.
2. Simple linear regression model.
3. Simple linear regression analysis.

1. The types of econometric variables. Difference Between Independent and Dependent Variable

Variable is a quantitative characteristic of the object, which can take different values in the process of the economic activity.

Types of quantitative variable:

- *Endogeneous* variables are internal variables of a model. Their values should be explain by an econometrics model;
- *Exogeneous* variables are external variables of a model. They should explain the value of the internal variable.

Another way to say, in the context of Statistical learning, there are two types of data:

- *Independent* variables: Data that can be controlled directly.
- *Dependent* variables: Data that cannot be controlled directly.

Dependent and independent variables are of particular interest

Difference Between Independent and Dependent Variable is shown in table 5.

Regression analysis is an important statistical method that allows us to examine the relationship between two or more variables in the dataset.

Regression analysis is used to:

- Predict the value of a dependent variable based on the value of at least one independent variable;
- Explain the impact of changes in an independent variable (x) on the dependent variable (y).

Table 5

Comparison Chart - Independent Variable Vs Dependent Variable

BASIS FOR COMPARISON	INDEPENDENT VARIABLE	DEPENDENT VARIABLE
Meaning	Independent Variable is one whose values are deliberately changed by the researcher in order to obtain a desired outcome.	Dependent Variable refers to a variable which changes its values in order to reciprocate change in the values of independent variable.
What is it?	The reason	Consequent
Relationship	Presumed cause	Observed effect
Values	Manipulated by the researcher.	Measured by the researcher.
Usually denoted by	x	y

2. Simple linear regression model

A **regression model** is a mathematical equation that describes the relationship between two or more variables. It is a transformation engine that helps us to express dependent variables as a function of independent variables.

Parameters are ingredients added to the model for estimating the output.

A **simple regression model** includes only two variables: one variable is independent and another one is dependent.

Example. Figure 4 is the graph of the equation $y = -1 + 2x$.

The most common and easiest way to display the relation between two variables x and y is a **scatter plot**. A scatter plot shows the direction and strength of a relationship between the variables.

Types of relations of two variables. They can be linear, nonlinear or curvilinear (e.g., parabola). In addition, there may be no relationship between them.

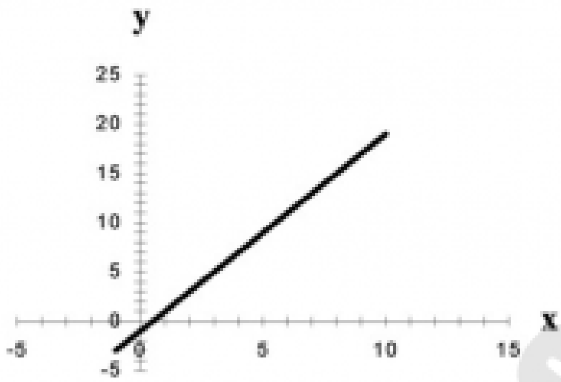


Fig. 4. The graph of the linear equation

Therefore simple linear regression estimates the relationship between a dependent variable and an independent variable using a straight line or curve.

"No dependence" means that we cannot describe the dependence between dependent and independent variables by the correct equation.

Linear implies the following: arranged in or extending along a straight or nearly straight line. Linear suggests that the relationship between dependent and independent variable can be **expressed in a straight line.**

A (simple) regression model that gives a straight-line relationship between two variables (*x* and *y*) is called a **linear regression model**.

$$y = a + bx \quad \text{or} \quad y = \beta_0 + \beta_1 x. \tag{5.1}$$

Linear regression is a manifestation of this equation.

Linear regression models are not perfect. It tries to approximate the relationship between dependent and independent variables in a straight line. Approximation leads to errors. Some errors can be reduced. Some errors are inherent in the nature of the problem. These errors cannot be eliminated. They are called as an **irreducible error**, the noise term in the true relationship that cannot fundamentally be reduced by any model.

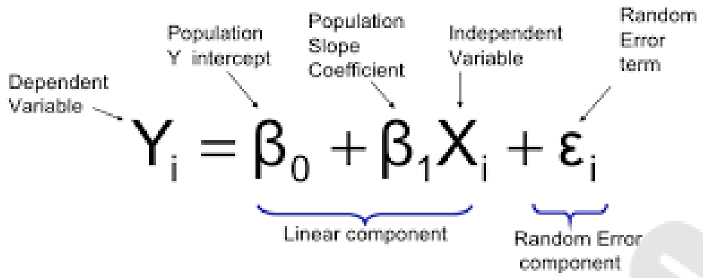
The same equation of a line can be re-written as:

$$y = a + bx + \varepsilon \quad \text{or} \quad y = \beta_0 + \beta_1 x + \varepsilon, \tag{5.2}$$

where a (β_0) and b (β_1) are the parameters; ε is the error term.

This part ($a + bx$) is the linear component of the regression equation and the other (ε) is the random error component.

3. Simple linear regression analysis



In the model $\hat{y} = \beta_0 + \beta_1 x$, β_0 and β_1, which are calculated using sample data, are called the estimates of β_0 and β_1, respectively.

- ***y*** is the dependent variable i.e. the variable that needs to be estimated and predicted.
- ***x*** is the independent variable i.e. the variable that is controllable. It is the input.
- β_1 is the slope. It determines what will be the angle of the line. It is the parameter denoted as β.
- β_0 is the intercept. A constant that determines the value of y when x is 0.

The Least Squares Method is the criteria for a “best” fit of linear equation.

The formula generates a line that represents the minimal distance between a predicted value for an independent variable and its actual value. The goal of linear regression is to minimize the total squared quantity for all these values. This is done through the method of least squares and can be calculated formulaically by methods such as the ordinary least squares (OLS) or iteratively by using machine-learning algorithms such as gradient descent. Before these calculations can be made, however, linear regression requires certain assumptions about data to be validated.

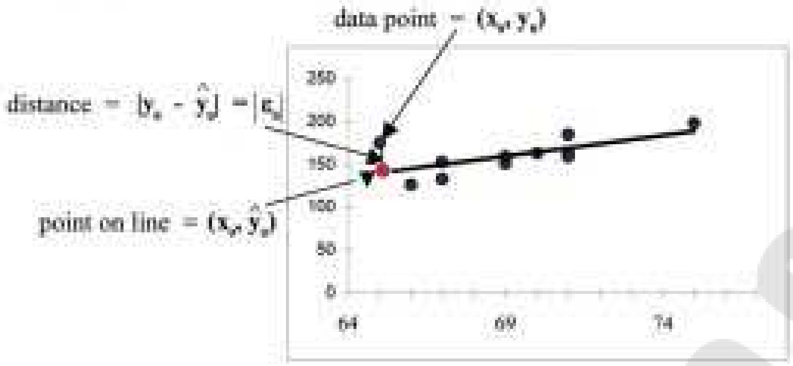


Fig. 5. The Least Squares Method

The term $y_i - \bar{y} = \varepsilon_i$ is called the "*error*" or *residual*. It is not an error in the sense of a mistake. The absolute value of a residual measures the vertical distance between the actual value of y and the estimated value of y. In other words, it measures the vertical distance between the actual data point and the predicted point on the line.

If the observed data point lies above the line, the residual is positive, and the line underestimates the actual data value for y. If the observed data point lies below the line, the residual is negative, and the line overestimates that actual data value for y.

β_0 and β_1 are obtained by finding the values of that minimize the sum of the squared differences between Y and:

$$min \sum (y_i - \hat{y}_i)^2 = min \sum \left(y_i - (\beta_0 + \beta_1 x_i)\right)^2. \qquad (5.3)$$

The error Sum of Squares, denoted SSE, is

$$SSE = \sum e^2 = \sum (y_i - \hat{y}_i)^2. \qquad (5.4)$$

The values of *a* (b_0) and *b* (b_1) that give the minimum SSE are called the *least square estimates* of *A* and *B*, and the regression line obtained with these estimates is called the *least square line*.

For the least squares regression line $\hat{y} = \beta_0 + \beta_1 x$,

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}, \qquad (5.5)$$

$$\hat{\beta}_0 = \bar{y} - \hat{b}_1 \bar{x}. \qquad (5.6)$$

The correlation coefficient r is a numerical measure of the strength of association between the independent variable x and the dependent variable y.

$$r = \frac{n \cdot \sum xy - (\sum x) \cdot (\sum y)}{\sqrt{[n \cdot \sum x^2 - (\sum x)^2] \cdot [n \cdot \sum y^2 - (\sum y)^2]}}, \tag{5.7}$$

where n is the number of data points.

If you suspect a linear relationship between x and y, then r can measure how strong the linear relationship is.

The value of r:

- The value of r is always between -1 and +1:
- The size of the correlation r indicates the strength of the linear relationship between x and y. Values of r close to -1 or to +1 indicate a stronger linear relationship between x and y.
- If $r = 0$ there is absolutely no linear relationship between x and y (no linear correlation).
- If $r = 1$, there is perfect positive correlation. If $r = 1$, there is perfect negative correlation. In both these cases, all of the original data points lie on a straight line.

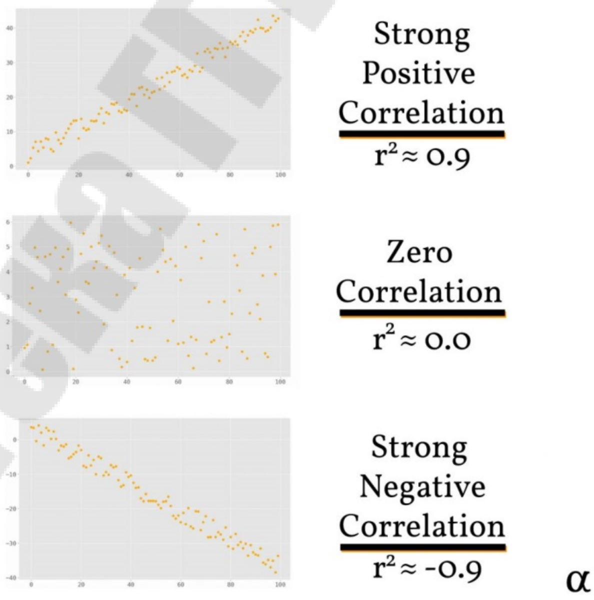


Fig. 6. A linear relationship cases

r^2 is called the coefficient of determination. r^2 is the square of the correlation coefficient , but is usually stated as a percent, rather than in decimal form. r^2 has an interpretation in the context of the data:

- r^2, when expressed as a percent, represents the percent of variation in the dependent variable y that can be explained by variation in the independent variable x using the regression (best fit) line.
- 1– r^2, when expressed as a percent, represents the percent of variation in y that is NOT explained by variation in x using the regression line. This can be seen as the scattering of the observed data points about the regression line.

TOPIC 6. MULTIPLE LINEAR REGRESSION MODELS

Plan:

1. The concept of multiple linear regression. Standardized regression coefficients and their interpretation.
2. Multiple regression model quality assessment.
3. Forecasting based on regression models.

1. The concept of multiple linear regression.Standardized regression coefficients and their interpretation

Multiple linear regression (MLR), also known simply as multiple regression, is a statistical technique that uses several explanatory variables to predict the outcome of a response variable.

Multiple linear regression is used to estimate the relationship between two or more independent variables and one dependent variable. You can use multiple linear regression when you want to know:

- How strong the relationship is between two or more independent variables and one dependent variable (e.g. how rainfall, temperature, and amount of fertilizer added affect crop growth).
- The value of the dependent variable at a certain value of the independent variables (e.g. the expected yield of a crop at certain levels of rainfall, temperature, and fertilizer addition).

Assumptions of multiple linear regression:

1. Multiple linear regression makes all of the same assumptions as simple linear regression:
2. Homogeneity of variance (homoscedasticity).

3. Independence of observations: the observations in the dataset were collected using statistically valid sampling methods, and there are no hidden relationships among variables.
4. In multiple linear regression, it is possible that some of the independent variables are actually correlated with one another, so it is important to check these before developing the regression model. If two independent variables are too highly correlated (r2 > ~0.6), then only one of them should be used in the regression model.
5. Normality.
6. Linearity.

In essence, multiple regression is the extension of ordinary least-squares (OLS) regression that involves more than one explanatory variable.

The formula for multiple linear regression is

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_n x_{in} + \varepsilon, \qquad (6.1)$$

where y = the predicted value of the dependent variable; β_0 = the y-intercept (value of y when all other parameters are set to 0); β_1 = the regression coefficient of the first independent variable (x_1) (a.k.a. the effect that increasing the value of the independent variable has on the predicted y value); … = do the same for however many independent variables you are testing; β_n = the regression coefficient of the last independent variable; ε = model error (a.k.a. how much variation there is in our estimate of y).

Multiple regression is an extension of simple linear regression in which more than one independent variable (x) is used to predict a single dependent variable (y). The predicted value of y is a linear transformation of the x variables such that the sum of squared deviations of the observed and predicted y is a minimum.

2. Multiple regression model quality assessment

To find the best-fit line for each independent variable, multiple linear regression calculates three things:

- The regression coefficients that lead to the smallest overall model error.
- The t statistic of the overall model.
- The associated p value (how likely it is that the t statistic would have occurred by chance if the null hypothesis of no relationship between the independent and dependent variables was true).

How well the equation fits the data is expressed by R^2, the "coefficient of multiple determination." The coefficient of multiple determination in multiple regression is similar to the coefficient of determination in simple linear regression, except in multiple regression there is more than one independent variable. The coefficient of multiple determination is the proportion of variation in the dependent variable that can be explained by the multiple regression model based on the independent variables.

This can range from 0 (for no relationship between *Y* and the *X* variables) to 1 (for a perfect fit, no difference between the observed and expected *Y* values). The *P* value is a function of the R^2, the number of observations, and the number of *X* variables.

Adjusted Coefficient of Multiple Determination

We use the adjusted coefficient of multiple determination, denoted adjusted R^2, which corrects the overestimation of the coefficient of multiple determination when new independent variables are added to the model. The adjusted coefficient of multiple determination is interpreted in the same way as the coefficient of multiple determination. The adjusted coefficient of multiple determination adjusts the value of R^2 to account for the number of independent variables in the model in order to avoid overestimating the impact of adding independent variables to the model.

The adjusted coefficient of multiple determination is calculated from the value of R^2:

$$adjusted\ R^2 = 1 - \frac{(n-1)\cdot(1-R^2)}{n-k-1}, \qquad (6.2)$$

where n is the number of observations; k is the number of independent variables.

Although we can find the value of the adjusted coefficient of multiple determination using the above formula, the value of the coefficient of multiple determination is found on the regression summary table.

3. Forecasting based on regression models

Forecasting models utilize historical and current information to provide a range of probable outcomes. The objective of a forecasting model is to extrapolate past and current trends with the help of various statistical and analytical tools to predict a future scenario. The results of such forecasting models form the basis of strategic decision making.

A forecasting model takes into account all the variables and possibilities that are associated with the subject to be forecasted. Such models are based on a number of assumptions, aggregations, and probabilities. Risk and uncertainty will, therefore, always underlie any forecasting model. The goal is only to forecast an outcome which would be closest to the real picture in order to minimize deviations from management expectations.

Using regression to make predictions doesn't necessarily involve predicting the future. Instead, you predict the mean of the dependent variable given specific values of the independent variable(s).

The general procedure for using regression to make good predictions is the following:

- Collect data for the relevant variables.
- Specify and assess your regression model.
- Use the actual or predicted values of the independent variables to calculate the values of the dependent variable using the resulting regression equation.

TOPIC 7. INVENTORY MANAGEMENT MODELS

Plan:

1. The definition of a stock. Types of stocks. Criteria for optimal inventory management.
2. The essence and characteristics of the main models of inventory management.

1. The definition of a stock. Types of stocks. Criteria for optimal inventory management.

Material cost shares a large portion on investment. Out of the total cost of production, material cost tunes up to 60-70%, sometimes even more than that. Hence, it calls in for a proper, timely and systematized Materials Management.

Inventory (**inventory stock**) – material values in the form of staple and semi-finished products at different stages of the production process and not applied in production at the moment, as well as finished products awaiting their enter into the production process or personal consumption.

Sustainable control of inventory stocks proposes the creation of such their level, which would ensure regular operation of the production process with minimum maintenance costs.

There are many reasons for creation of material stocks and inventories in companies; however, what unites us is striving of production activity subjects towards the economic security. With this it should be noted that the cost of inventory stocks creation and ambiguity of terms of marketing conditions do not stimulate the increasing importance of expensive backup network "security” in the eyes of company’s management, because they [marketing and inventory stocks] objectively contradict to the increase of production effectiveness.

One of the main incentives for the creation of inventory stocks is the cost of their negative level (deficit). In case of inventory deficit, there are three types of possible expenditures, which are listed below according to their increasing negative influence:

- expenditures due to non-performance of an order (delay with sending the ordered goods) – additional expenditures for promotion and dispatch of goods of the order that cannot be performed on the means of the existing inventory and material stocks.
- expenditures due to marketing losses in cases when the regular customer applies for certain purchase in some other company (such expenditures are measured by the index, showing the loss of revenue because of unrealized bargain);
- expenditures due to loss of customer in cases when absence of inventory stocks results not only in waste of bargain of some sort, but also in the fact that customer starts looking for an alternative sources of supply (such expenditures are measured in terms of total revenue, which could have been gained from realization of all potential bargains between the customer and company).

The inventory problem reduces to devising an **inventory policy** that answers two questions:

1. *How much* to order?
2. *When* to order?

The basis of the inventory model is the following generic cost function:

$$\begin{pmatrix}\text{Total}\\ \text{inventory}\\ \text{cost}\end{pmatrix} = \begin{pmatrix}\text{Purchasing}\\ \text{cost}\end{pmatrix} + \begin{pmatrix}\text{Setup}\\ \text{cost}\end{pmatrix} + \begin{pmatrix}\text{Holding}\\ \text{cost}\end{pmatrix} + \begin{pmatrix}\text{Shortage}\\ \text{cost}\end{pmatrix}$$

The described costs are conflicting, in the sense that an increase in one may result in the reduction of another (e.g., more frequent ordering results in higher setup cost but lower inventory holding cost). The purpose of the minimization of the total inventory cost function is to balance these conflicting costs.

To make a decision, ***methods and models of inventory control theory*** are used. Based on this field's obtained results, practical algorithms and computer programmers for inventory management in systems of production and logistics are developed and implemented.

Creation of inventory is always associated with expenditures. First, inventory stocks is actually blocked financial means; second, maintenance costs for specially equipped premises (warehouses) and payment to the personnel; third, the risk of damage and theft of stocks, which results in extra expenditures as well. At the same time, the absence of invention stocks can also result in expenditures, which appear in the form of different losses: because of production downtime in case of goods delayed delivery; losses from buying up at high rates and transportation of small consignments; losses because the goods are out of stock when they are in demand.

2. The essence and characteristics of the main models of inventory management.

All models are classified into two major types:

1. Deterministic Models,
2. Probabilistic Models.

In brief, the *deterministic models* are built on the assumption that there is no uncertainty associated with demand and replenishment of inventories. On the contrary, the *probabilistic models* take cognizance of the fact that there is always some degree of uncertainty associated with the demand pattern and lead time of inventories.

On the basis of objectives of inventory management mentioned earlier, the major problems faced by inventory management are to determine:

- Level of inventory at which new purchase order should be placed,
- Quantity of material which should be purchased by each purchase order, and
- Type of control required for each type of inventory.

Various techniques of inventory management are adopted to find the answers of these problems. The most well-known deterministic model is the Wilson model/

Wilson's model. It is also called the Optimal Order model or **EOQ model**. It is based on mathematical formulas to define the most indicated quantities of orders that must be made in the company to make the investment of assets more efficient.

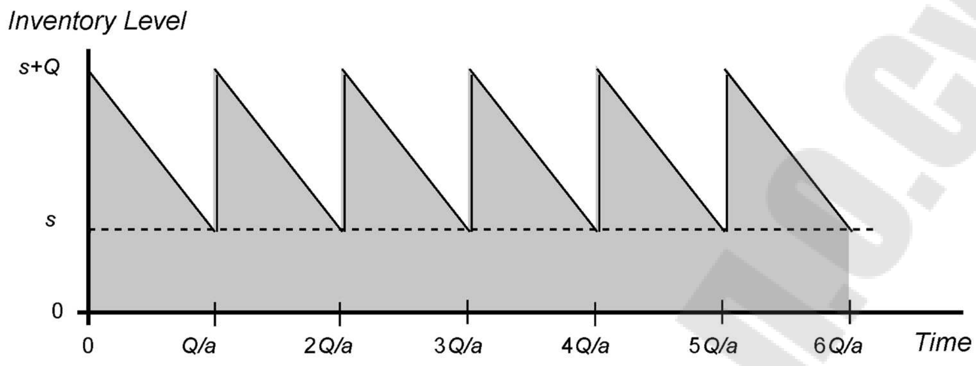


Fig. 7. Inventory management model without uncertainty

The factors that are important in making decisions related to inventories:

- ***Ordering cost (c(z))***: This is the cost of placing an order to an outside supplier or releasing a production order to a manufacturing shop. The amount ordered is *z* and the function *c(z)* is often nonlinear. The dimension of ordering cost is ($).
- ***Setup cost (K)***: A common assumption is that the ordering cost consists of a fixed cost, that is independent of the amount ordered, and a variable cost that depends on the amount ordered. The fixed cost is called the setup cost and given in ($).
- ***Product cost (c)***: This is the unit cost of purchasing the product as part of an order. If the cost is independent of the amount ordered, the total cost is *cz*, where *c* is the unit cost and *z* is the amount ordered. Alternatively, the product cost may be a decreasing function of the amount ordered. ($/unit)
- ***Holding cost (h)***: This is the cost of holding an item in inventory for some given unit of time. It usually includes the lost investment income caused by having the asset tied up in inventory. This is not a real cash flow, but it is an important component of the cost of inventory. If *c* is the unit cost of the product, this component of the cost is $c\alpha$, where is α the discount or interest rate. The holding cost may also include the cost of storage, insurance, and other factors that are proportional to the amount stored in inventory. ($/unit-time)

- ***Shortage cost (p)***: When a customer seeks the product and finds the inventory empty, the demand can either go unfulfilled or be satisfied later when the product becomes available. The former case is called a lost sale, and the latter is called a backorder. The total backorder cost is assumed to be proportional to the number of units backordered and the time the customer must wait. The constant of proportionality is p, the per unit backorder cost per unit of time. ($/unit-time)
- ***Demand rate (a)***: This is the constant rate at which the product is withdrawn from inventory. (units / time)
- ***Lot Size (Q)***: This is the fixed quantity received at each inventory replenishment. (units)
- ***Order level (S)***: The maximum level reached by the inventory is the order level. When backorders are not allowed, this quantity is the same as Q. When backorders are allowed, it is less than Q. (units)
- ***Cycle time (τ)***: The time between consecutive inventory replenishments is the cycle time. For the models $\tau = Q/a$. (time)
- ***Cost per time (T)***: This is the total of all costs related to the inventory system that are affected by the decision under consideration. ($/time)
- ***Optimal Quantities (Q*, S*, τ*, T*)***: The quantities defined above that maximize profit or minimize cost for a given model are the optimal solution.

EOQ can be explained with the help of following diagram:

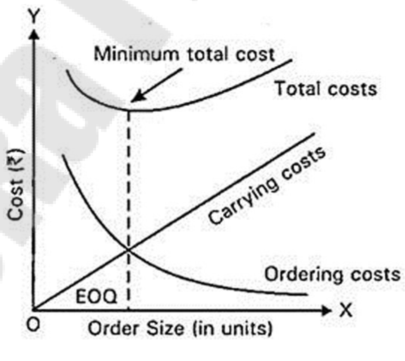


Fig. 8. EOQ model illustration

The total cost expressed per unit time is

Cost/unit time = Setup cost + Product cost + Holding cost

$$T = \frac{aK}{Q} + ac + \frac{hQ}{2}, \quad (7.1)$$

$\frac{a}{Q}$ is the number of orders per unit time. The factor $\frac{Q}{2}$ is the average inventory level.

From here

$$Q^* = \sqrt{\frac{2aK}{h}}, \tag{7.2}$$

$$\tau^* = \frac{Q^*}{a}. \tag{7.3}$$

Substituting the optimal lot size into the total cost expression, Eq. (7.1), and preserving the breakdown between the cost components we see that

$$T^* = ac + \sqrt{2ahK}. \tag{7.4}$$

For this model, the optimal policy does not depend on the unit product cost. The optimal lot size increases with increasing setup cost and flow rate and decreases with increasing holding cost.

TOPIC 8. NETWORK PLANNING AND MANAGEMENT

Plan:

1. Basic concepts of network planning. Methods of parameter calculations network schedule.
2. Concept and criteria of optimization of network graphs.

1. Basic concepts of network planning. Methods of parameter calculations network schedule

The success of a project's realization depends heavily on the effectiveness of the planning phase.

Network planning is a common term for methods where projects are studied as a series of interrelated activities to plan, manage, and control projects.

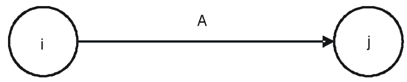
The most common network planning tools are the CPM and PERT.

The critical path method (CPM) is a project scheduling algorithm used to calculate the total time needed to complete a project.

In the critical path method, the critical activities of a program or a project are identified. These are the activities that have a direct impact on the completion date of the project.

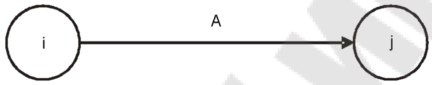
CPM/PERT networks contains two major components:

Activity: An activity represents an action and consumption of resources (time, money, energy) required to complete a portion of a project. Activity is represented by an arrow,



A is called activity

Event: An event (or node) will always occur at the beginning and end of an activity. The event has no resources and is represented by a circle. The *i*th event and *j*th event are the *tail event* and *head event* respectively,

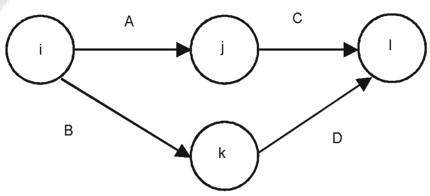


Tail event Head event

One or more activities can start and end simultaneously at an event

Merge event Burst event

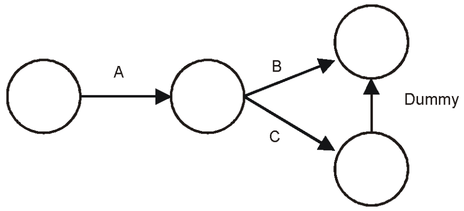
Activities performed before given events are known as preceding activities, and activities performed after a given event are known as succeeding activities.



Activities A and B precede activities C and D respectively.

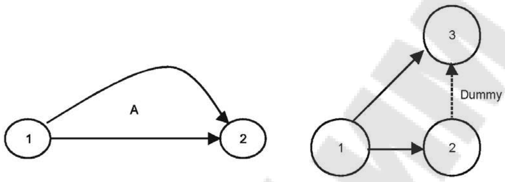
Dummy Activity

An imaginary activity which does not consume any resource and time is called a dummy activity. Dummy activities are simply used to represent a connection between events in order to maintain a logic in the network. It is represented by a dotted line in a network.



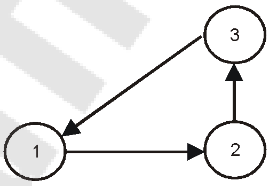
Errors to be avoided in Constructing a Network

1. Two activities starting from a tail event must not have a same end event. To ensure this, it is absolutely necessary to introduce a dummy activity.
2.



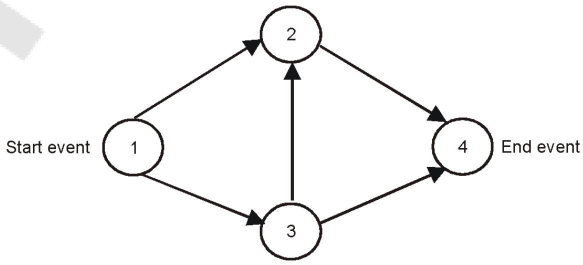
Incorrect Correct

3. Looping error should not be formed in a network, as it represents performance of activities Notes repeatedly in a cyclic manner,
4.



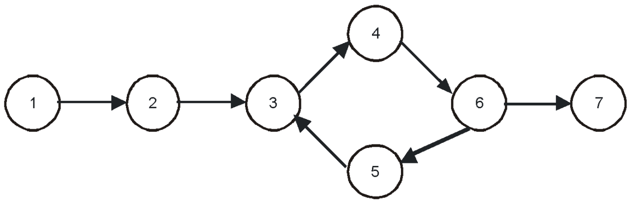
Incorrect

5. In a network, there should be only one start event and one ending event as shown below,



6. The direction of arrows should flow from left to right avoiding mixing of direction as shown below

7.



Incorrect

Rules in Constructing a Network

1. No single activity can be represented more than once in a network. The length of an arrow has no significance.
2. The event numbered 1 is the start event and an event with highest number is the end event.
3. Before an activity can be undertaken, all activities preceding it must be completed. That is, the activities must follow a logical sequence (or interrelationship) between activities.
4. In assigning numbers to events, there should not be any duplication of event numbers in a network.
5. Dummy activities must be used only if it is necessary to reduce the complexity of a network.

   A network should have only one start event and one end event.

Some conventions of network diagram is shown in Figure 8 (a), (b), (c), (d) below:

Key elements of Critical Path Method

- *The project activity list* is a comprehensive list of all the tasks and work items required to complete the project. The project activity list is the foundation for the project network diagram.
- *The project network diagram* is a graphical representation of the project schedule that shows the interdependencies between tasks. It calculates the critical path and determines the float/slack for each task.

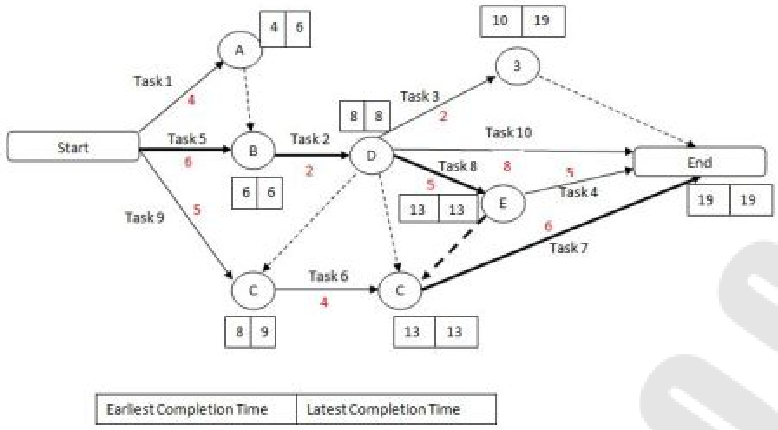
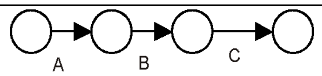
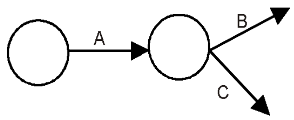
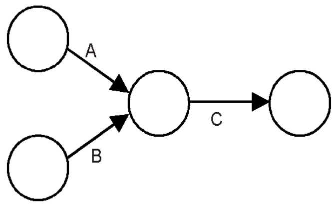
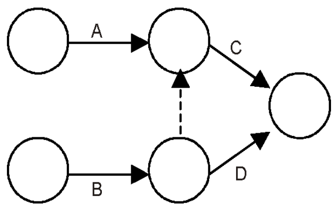


Fig. 8. The project network diagram

	(a) Activity B can be performed only after completing activity A, and activity C can be performed only after completing activity B.
	(b) Activities B and C can start simultaneously only after completing A.
	(c) Activities A and B must be completed before start of activity C.
	(d) Activity C must start only after completing activities A and B. But activity D can start after completion of activity B.

The **critical path** for any network is the longest path through the entire network. Since all activities must be completed to complete the entire project, the length of the critical path is also the shortest time allowable for completion of the project. Thus if the project is to be completed in that shortest time, all activities on the critical path must be started as soon as possible. These activities are called critical activities. If the project has to be completed ahead of the schedule, then the time required for at least one of the critical activity must be reduced. Further, any delay in completing the critical activities will increase the project duration.

The activity, which does not lie on the critical path, is called non-critical activity. These non-critical activities may have some slack time.

The slack is the amount of time by which the start of an activity may be delayed without affecting the overall completion time of the project. But a critical activity has no slack. To reduce the overall project time, it would require more resources (at extra cost) to reduce the time taken by the critical activities to complete.

- *Project duration* is the time required to complete a project. Project duration is a critical factor in CPM as it directly impacts the budget and overall success.
- *Resource constraints,* such as limited availability of personnel, equipment, or materials, can impact the duration of tasks and ultimately affect the overall project timeline.
- *Early Start and Early Finish* are a task's earliest possible start and finish times based on the project schedule and the interdependencies between tasks. Early Start and Early Finish provide project managers with a baseline for the project schedule and are used to determine the critical path.
- *Late Start and Late Finish* are a task's latest possible start and finish times based on the project schedule and the interdependencies between tasks. Late Start and Late Finish provide project managers with an understanding of the flexibility in the project schedule and are used to determine float/slack.
- *Float/Slack (reserve)* refers to the time a task can be delayed without affecting the overall project timeline. It is the difference between a task's Late Start and Early Start or a task's Late Finish and Early Finish.

Key Steps in Critical Path Method

1. List all tasks and estimate their duration
2. Determine task dependencies
3. Create the project network diagram
4. Calculate the earliest start and finish times
5. Calculate the latest start and finish times
6. Identify critical path tasks
7. Update the project network diagram

Time Estimates:

Forward Pass Computations (to calculate Earliest, Time T_E)

Procedure

Step 1: Begin from the start event and move towards the end event.

Step 2: Put $T_E = 0$ for the start event.

Step 3: Go to the next event (i.e. node 2) if there is an incoming activity for event 2, add calculate T_E of previous event (i.e. event 1) and activity time. Note: If there are more than one incoming activities, calculate T_E for all incoming activities and take the maximum value. This value is the T_E for event 2.

Step 4: Repeat the same procedure from step 3 till the end event.

Backward Pass Computations (to calculate Latest Time T_L)

Procedure

Step 1: Begin from end event and move towards the start event. Assume that the direction of arrows is reversed.

Step 2: Latest Time *TL* for the last event is the earliest time. T_E of the last event.

Step 3: Go to the next event, if there is an incoming activity, subtract the value of T_L of previous event from the activity duration time. The arrived value is T_L for that event. If there are more than one incoming activities, take the minimum T_E value.

Step 4: Repeat the same procedure from step 2 till the start event.

Determination of Float and Slack Times

The float of an activity is the amount of time available by which it is possible to delay its completion time without extending the overall project completion time.

For an activity

$i = j$, let

t_{ij} = duration of activity

T_E = earliest expected time

T_L = latest allowable time

ES_{ij} = earliest start time of the activity

EF_{ij} = earliest finish time of the activity

LS_{ij} = latest start time of the activity

LF_{ij} = latest finish time of the activity

Total float TF_{ij}: The total float of an activity is the difference between the latest start time and the earliest start time of that activity.

$$TF_{ij} = LS_{ij} - ES_{ij}, \tag{8.1}$$

or

$$TF_{ij} = (T_L - T_E) - t_{ij}. \tag{8.2}$$

Free Float FF_{ij}: The time by which the completion of an activity can be delayed from its earliest finish time without affecting the earliest start time of the succeeding activity is called free float:

$$FF_{ij} = (E_j – E_i) – t_{ij}, \quad (8.3)$$
$$FF_{ij} = Total\ float – Head\ event\ slack$$

Independent Float IF_{ij}: The amount of time by which the start of an activity can be delayed without affecting the earliest start time of any immediately following activities, assuming that the preceding activity has finished at its latest finish time:

$$IF_{ij} = (E_j – L_i) – t_{ij}, \quad (8.4)$$
$$IF_{ij} = Free\ float – Tail\ event\ slack$$

where tail event slack = L_i - E_i.

Critical Path: After determining the earliest and the latest scheduled times for various activities, the minimum time required to complete the project is calculated. In a network, among various paths, the longest path which determines the total time duration of the project is called the critical path. The following conditions must be satisfied in locating the critical path of a network.

An activity is said to be critical only if both the conditions are satisfied:

$$1.\ T_L – T_E = 0, \quad (8.5)$$
$$2.\ T_{Lj} – t_{ij} – T_{Ej} = 0. \quad (8.6)$$

2. Concept and criteria of optimization of network graphs

The initially developed network model is usually not the best in terms of work completion and resource usage. Therefore, the original network model is being optimized.

Network graphs optimize:

a) by timing;

b) by the resources used;

c) by cost.

Resource allocation is the process of assigning and scheduling the available resources, such as people, materials, equipment, and budget, to the tasks and activities of a project. Resource allocation aims to balance the demand and supply of resources, and to ensure that the project objectives are met within the scope, time, and quality constraints.

When optimizing the use of labor resources, most often network work tends to be organized in such a way that:

• the number of simultaneously employed performers was minimal;

• Equalize the need for human resources over the duration of the project.

The essence of optimizing the loading of network models according to the criterion of "minimum performers" is as follows: it is necessary to organize the work in such a way that the number of simultaneously working performers is minimal. To carry out such types of optimization, it is necessary to build and analyze the binding schedule and the loading schedule.

Table 6

Initial data for load optimization

Work code	Duration of work	Number of performers
(1,2)	4	6
(1,3)	3	1
(1,4)	5	5
(2,5)	7	3
(2,6)	10	1
(3,6)	8	8
(4,6)	12	4
(4,7)	9	2
(5,8)	8	6
(6,8)	10	1
(7,8)	11	3

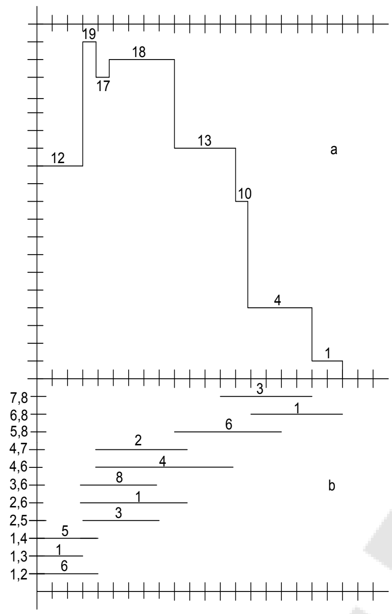


Fig. 9. Loading graphs (a) and bindings (b) before optimization

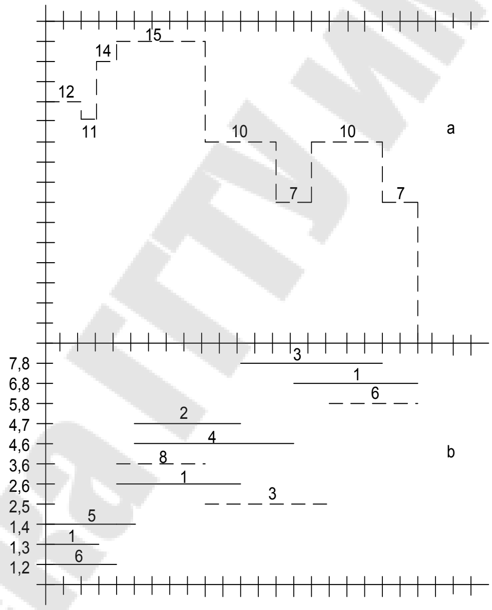


Fig. 10. Loading graphs (a) and bindings (b) after optimization

As a result of optimization, the maximum load of the network model decreased from 19 to 15 people, which was the goal of the optimization.

The list of recommended literature

1. Borodich, S.A. Econometrics: tutorial / S.A.Borodich. – 3th ed., rev. – Mn.: New knowledge, 2006. – 408 p.
2. Cramer, N.Sh. Econometrics: a textbook / N.Sh. Cramer, B.A. Putko; under the ed. N.Sh. Cramer. – 2th ed. – M.: UNITI-DANA, 2008. – 312 p.
3. Magnus, Ya.R. Econometrics: The initial course / Ya. R. Magnus, P. K. Katyshev, A. A. Peresetsky. - Moscow : Delo, 1997. - 248 p.
4. Mardas, A. N. Econometrica : textbook / A. N. Mardas. - St. Petersburg : Peter, 2001. - 144 p.
5. Econometrics and economic-mathematical methods and models: tutorial for universities / G.O. Chitaya and [oth.]; under the ed. G.O. Chitaya, S.F. Miksyuk. –Mn.: BSEU, 2018. – 510 p.
6. Belko, I. V. Econometrica: practicum / I. V. Belko, E. A. Krishtapovich. - Minsk : Grevtsov Publishing House, 2011. – 220 p.
7. Verbeek, M. A guide to modern econometrics. - 2nd ed. – Wiley, 2004. – 446 p. [Electronic resource.] – Access mode: http://bookre.org/reader?file=1196330
8. Dougherty Ch. Introduction to econometrics. – 4th edition. – Oxford University Press, 2011. – 573 p. [Electronic resource]. – Access mode: http://bookre.org/reader?file=726521

Vinnik Olga Grigorievna

ECONOMETRICS

Textbook
for specialty 6-05-0311-02 “Economics and management”
for full-time students

Подписано к размещению в электронную библиотеку
ГГТУ им. П. О. Сухого в качестве электронного
учебно-методического документа 14.11.24.
Рег. № 10E.
http://www.gstu.by