INVESTIGAÇÃO OPERACIONAL

Junho 1987
Número 1
Volume 7

Publicação Científica da
apdio
Associação Portuguesa para o Desenvolvimento da Investigação Operacional.

INVESTIGAÇÃO OPERACIONAL

Propriedade:

APDIO — Associação Portuguesa para o Desenvolvimento da Investigação Operacional

ESTATUTO EDITORIAL

«Investigação Operacional», órgão oficial da APDIO cobre uma larga gama de assuntos reflectindo assim a grande diversidade de profissões e interesses dos sócios da Associação, bem como as muitas áreas de aplicação da I. O. O seu objectivo primordial é promover a aplicação do método e técnicas da I.O. aos problemas da Sociedade Portuguesa.
A publicação acolhe contribuições nos campos da metodologia, técnicas, e áreas de aplicação e software de I. O. sendo no entanto dada prioridade a bons casos de estudo de carácter eminentemente prático.

Distribuição gratuita aos sócios da APDIO

INVESTIGAÇÃO OPERACIONAL

volume 7 - nº 1 - Junho 1987 Publicação semestral

A Revista "INVESTIGAÇÃO OPERACIONAL" está registada na Secretaria de Estado da Comunicação Social sob o nº108335.
Esta Revista é distribuida gratuitamente aos sócios da APDIO. As informações sobre inscrições na Associação, assim como a correspondência para a Revista devem ser enviadas para a sede da APDIO - Associação Portuguesa para o Desenvolvimento da Investigação Operacional - CESUR, Instituto Superior Técnico, Av. Rovisco Pais, 1000 Lisboa.

Este Volume foi subsidiado por :

Instituto Nacional de Investigação Científica (INIC)

Junta Nacional de Investigação Científica e Tecnológica (JNICT)

Fundação Calouste Gulbenkian

Para efeitos de dactilografia e composição, foram utilizados equipamentos gentilmente postos à disposição pelo CEAUL (DEIOC- Faculdade de Ciências de Lisboa).

Assinatura : 3000$00

Investigação Operacional, Vol.7 nº1 (Junho/1987) 3 - 10

OPERATIONS RESEARCH IN BANKING. PAST SUCCESSES & FUTURE OPPORTUNITIES.

Frederick J. Rigdway

Bank of Ireland

SCOPE OF PAPER

I intend to cover in this paper

- the Development of O.R. in Banking from the 1960's to the 1980's.
- the current areas of active O.R. work.
- future Developments of O.R. in Banking.

1. A Personal Perspective

I commenced my career in Banking in 1967, when I was recruited by the Bank of Ireland in Dublin to set up its new Operations Research function. The Bank of Ireland, like many other Banks in Western Europe, had at that time completed a major reorganisation. This was the time when smaller Banks were merging into large "Super Banks". In Britain, we saw the four major Clearing Banks emerge as the Nat West, Barclays, Midland and Lloyds. The merger of a number of smaller Banks into a "Super Bank" gave rise to tremendous opportunities for the application of Operations Research in Banking. I was very privileged to participate in this very active and interesting era. It meant that many old ideas were challenged and many new ways of looking at Bank Management were considered and implemented. In this Paper, I will be setting out for you, the major areas in which Operations Research contributed to banking, the current areas of active research, and my views on the future development of Operations Research in Banking.

As I pointed out, the Bank of Ireland was not alone in its growth through acquisition and merger. Many major European Banks went through the same process at that time. It was, therefore, of great benefit to Operations Research workers to exchange ideas on mutually shared problems. A group of O.R. workers in Banking formed, in the late 60's, a European Working Group for O.R. This Group has held an annual meeting each year since its inception. At these meetings there has always been a very frank and open exchange of ideas. As a result, the developments I will be describing to you reflect not just my own personal experience, but the experience gained from attending these meetings both as a speaker and a Participant.

2. The Role of Operations Research in Banking

I will identify two distinct roles for Operations Research in Banking:

(a) Strategic O.R.

(b) Workplace O.R.

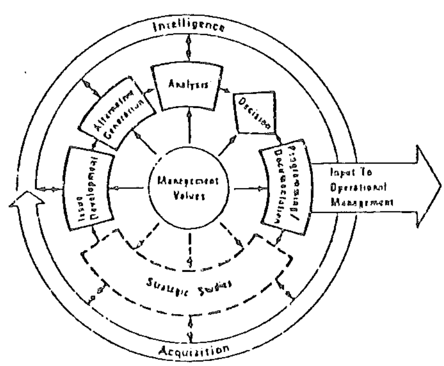
By Strategic O.R. I mean the applications of O.R. methods to major Bank wide issues of a long term and fundamental nature.

By Workplace O.R. I mean the application of O.R. techniques to tactical problems for example scheduling work in the Bank's Computer Centre or the optimal distribution of cash to Branches. Both types of O.R. work have contributed effectively to the development of the Banking industry.

However, it is my opinion that the Strategic O.R. area has made a remarkable contribution to European Banking development.

EXHIBITION 1

THE STRATEGIC MANAGEMENT CYCLE



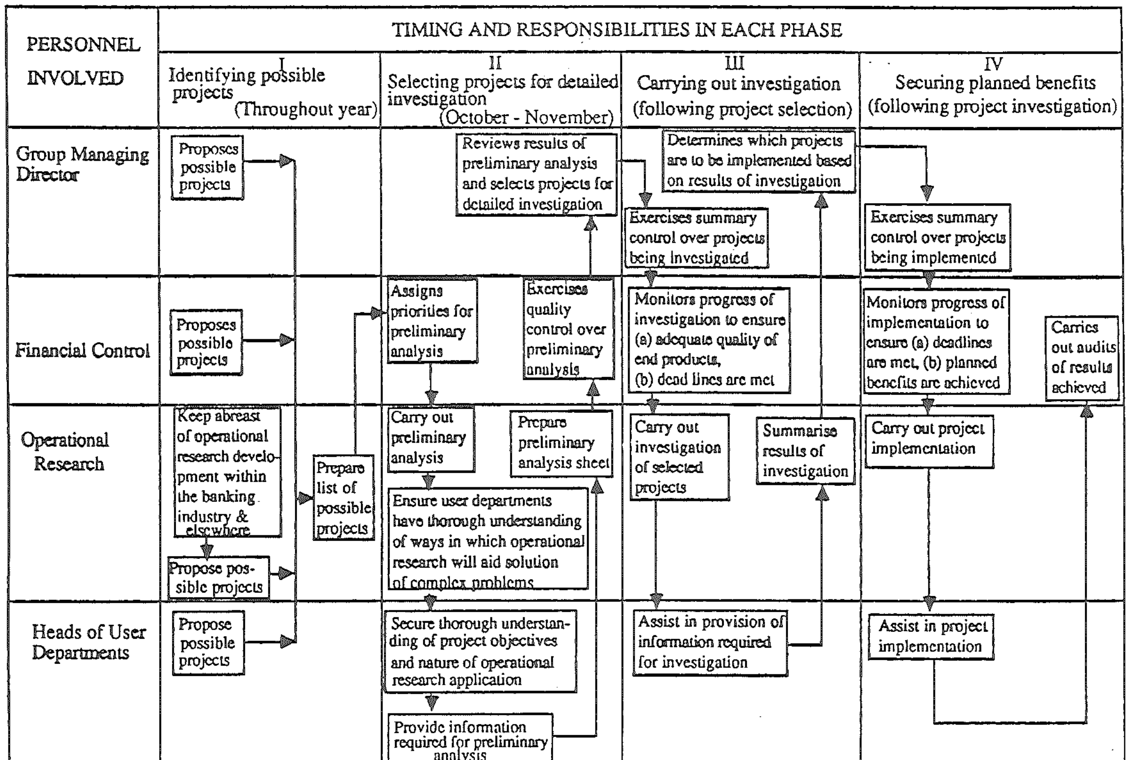
I believe that Strategic O.R. is one of the most cost effective ways of developing strategy. Exhibit 1 shows a typical approach to the development of Corporate strategy. To support this activity the O.R. approach of model building and sampling data rather than building data bases is an exceedingly effective way of developing and quantifying strategic alternatives. Building large data bases is expensive, and usually by the time they are complete the issues for which they were to suply the information have probably already been solved. With the O.R. approach, models can be built quickly, the data for them being obtained by statistical sampling and thus Bank Management can have quantified alternatives on which to develop their strategy.

3. Location of Operations Research in a Bank

If a Bank is to make the most effective use of its O.R. skills, experience indicates that the O.R. group should be located in a Corporate Planning or an Economic Planning Department. These are usually located at a high level in the organisation's hierarchy. This organisational choice has been taken by Banks in Ireland, United Kingdom, The Netherlands, France, Scandinavia and Italy, where it has been proved conclusively that

RESPONSABILITIES OF GROUP PERSONNEL IN PLANNING AND CONTROLLING THE OPERATIONAL RESEARCH EFFORT

EXHIBITION 2



O.R. can tackle the big problems in Banking. Many of the original group of O.R. workers in these Banks now occupy very senior positions in Line Management. In one of the Banks (one of the largest in the World) there is a Management policy of moving young high-flyers through the O.R. department as a necessary step in their career development. During their time in the O.R. Department they have the opportunity of seeing major issues being developed and can see how problems affect many departments in a Bank.

4. How to Make O.R. Succeed in Bank's - Management Role

To make O.R. effectively contribute to a Bank's profit improvement programme, Management must:

4.1 Locate the O.R. group where it can be most effective and as I have already suggested this should be at a high level in the Corporate Planning or Economic Planning Departments.

4.2 In the selection of personnel for certain functions, the technical competence of individuals must be matched by their ability to communicate effectively at all levels. I would suggest that in selecting O.R. people, a good criteria would be their potential for development to senior positions long term in the Bank.

4.3 The O.R. group should be kept small but there is a minimum number below which effective work cannot be done. It is my opinion that the minimum size should be 3 specialists and 3 technical support staff.

4.4 There must be effective Management Planning and control of the O.R. function. Exhibit 2 shows a suggested system of planning and control. You will see that this is based on an annual cycle. I consider that the most important part of this cycle is the project selection phase. Exhibit 3 shows this in more detail.

Management and O.R. workers, alike, must place a very high priority on the work of selecting O.R. projects. Great emphasis must be placed on identifying the profit improvement opportunities associated with each of the projects. Management must be convinced ab initio that the benefits are realisable and as the Flow Chart shows (Exhibit 2), progress in securing these benefits should be carefully monitored. Profit improvement must be the clear goal of the O.R. group. If this goal is lost sight of, or Management does not exercise proper control, there is a danger that Operations Research can stray into doing work which, while being intellectually interesting, does not contribute to the well-being of the Bank.

5. How to Maker O.R. Succeed in Banking - The Role of the O.R. Group

As in the case of Management, the most important task for the O.R. group is the selection of projects, particularly when a new O.R. group is set up. The first projects selected will be the key to their ongoing success. It is my experience that, inevitably there will be good projects, with significant profit potential for a new group, provided it is properly located near the centre of power in the Bank. Initially, the group should concentrate on a single project. A significant proportion of the group's time should be spent on communicating progress to Management. Results, however good, are of little value if they are not effectively communicated. All major communications media should be used. I strongly believe that any written reports must be supported by a verbal presentation with very high quality visual aids.

We can learn here from International Management Consultants firms who have to earn their living by effectively communicating their findings to top Management. In these companies you will be well aware of the time and effort spent in preparing reports and visual aids and in rehearsing the presentations for Management. For young O.R. workers coming into Banking organisations with a good academic record, communication is probably the area of their skills which need most development. Both the Bank and the individual should be aware of this and a special effort made to develop communications skills.

6. How O.R. has Developed in the Main Application Areas

6.1 Assets/Liabilities Management

Linear Programming has long been one of the most widely taught O.R. techniques in Universities. It was natural, therefore, for O.R. workers coming into banking to look for applications of this powerful technique in banking. Asset/Liability Management at a Corporate level offers the single most important profit improvement opportunity in any bank. Very small increases in the margin achieved on funds can give rise to very substantial profit improvements. The application of L.P. to Asset/Liability Management became a major work area in newly established O.R. groups during the late 60's and early 70's. Leading workers in this field were Cohan & Hammer in the United States working at Bankers Trust. Unfortunately, "the grand" approach did not prove successful. I believe the main problem was the high level of uncertainty in the prediction of Interest Rates and the repayment pattern on Loans. In Europe also a lot of work was devoted to the application of Linear Programming to this application again, with in my opinion, indifferent results. However, this work was not completely without value because in preparing the information for the L.P. model many interesting profit improvement opportunities were identified in individual assets/liabilities. What developed was the modeling of the Bank's balance sheet and profit/loss statement on a simulation model approach. This gave Management an opportunity to see the consequences of various scenarios and the application of strategic decisions to these scenarios. This approach has been very fruitful and from it many projects in relation to particular assets/liabilities have been generated; in particular development of new Interest Rate structures for both Deposits and Loans, and the use of statistical sampling and simulation to predict the effects of these changes on the Bank's profit stream.

6.2 Interest Rate Forecasting

Also arising from the original L.P. work, it became clear that the prediction of Interest Rates was a major problem area. Research work was undertaken into Interesting Rate prediction - for example by the application of the Delphi approach. This was a technique developed by the Rand Corporation in California, initially for military applications. The Delphi method is a way of combining expert opinion on future events. In Bank's there are a number of people who, from their experience of Markets and their Economic specialised knowledge, will have expert opinion on how Interest Rates will move. What Bank of Ireland did and also a number of other Banks, was to use the Delphi method as developed by the Rand Corporation to combine these expert opinions to produce a group forecast of Interest Rates. A combination of advance time series analysis for forecasting movements on balance sheet items such as Loans/Deposits and use of the Delphi method to forecast Interest Rates has offered the Banks a very tangible benefit in the field of Asset/Liability Management.

6.3 Cash Management

The field of asset/liability Management would not be complete without mention of Branch Cash Management. Commercial Banks normally hold a large amount of Cash in their Branches and in transit between Branches and their Central Treasury. In most countries cash in Branches does not earn any interest. However, if surplus cash can be lodged at a Central Bank depot, it then earns interest. Clearly this is a classical application of O.R. techniques of inventory and distribution Management. Many interesting applications have been made in this area and very large profit improvements have been made for their Banks by the O.R. groups. In this area the use of Micro-computers has been successful.

6.4 Manpower Planning

In examining a Bank's profit/loss statement, one of the major components in that statement is Staff Costs. Banks in general tend to be stable employers taking on Staff at school-leaving age and retaining them until retirement. Investment in a new employee represents a committment over a lifetime.

6.4.1-It is very important, therefore, for the Bank to know how many Staff it needs to recruit each year, because of the implications of this recruitment for the long term pay structure in the Bank.

6.4.2-The Bank also needs to know how the pattern of staff by age and grade will change over future years. Will there be a shortage of Management potential due to a large number of senior Manager retiring, and only inadequately trained junior Managers being available? Will there be a serious block in promotion for a large number of talented people competing for a relatively small number of Managerial vacancies?

6.4.3-What will the Staff Costs be on into the future as the proposed recruitment policy is carried out over the planning horizon?

These are some of the interesting problems that Management have to face in the area of Manpower Planning and Staff Costs. This is an ideal application for O.R. modelling and has been one of the most successful applications carried out in European Banks. It has all the ingredients of a good O.R. application, model building, **stocastic** processors, forecasting models. The basic modelling approach is now well known and documented, and its application to the individual Bank offers a challenging and interesting project for any O.R. group. What is particularly important about Manpower Planning is that it should be integrated into the normal planning cycle of the Bank. Recently a new dimension for Manpower planning has been opened up with the advent of Bank Automation. Modelling the impact of Staff requirements as a result of the introduction of Lobby Banking, Automatic Teller Machines, Home Banking and Self Service all give rise to further worthwhile uses for Manpower Planning models. Without such models the full impact of large scale technology programmes will be very difficult to assess.

6.5 Statistical Cost Analysis

If one makes a crude breakdown of the Costs of providing services, one of the most costly services provided by a retail Bank network is its Money Transmission service. This service along with Deposit taking and Lending normally constitute the major proportion of a Bank's activities. Because Costs arise at different points in a Bank network and a costumer can also transact business at any one of many outlets-the costing of individual transactions, services and products is very difficult. This costing problem has proved to be a most successful area for O.R. The technique used here is Statistical Cost Analysis and using this technique Banks have been able to allocate costs to their individual products and services thus forming a basis of individual customer profitability analysis. Once the fundamental approach has been established and sampling methods effected, an O.R. group can provide invaluable service in specialised costing applications e.g. in the area of new E.D.P. applications, changes in systems and procedures. A proper knowledge of costing is the key to a whole range of strategic issues particularly in the field of Marketing, Pricing and Technology development.

6.6 Corporate Planning Model

The applications I have so far listed all form building blocks for developing a Bank's Corporate Planning Model. Such models form an important part of the Strategic Management cycle as described in Exhibit 1. Large scale simulation models drawing information from Asset/Liabilities Models, Manpower Planning Models, Statistical Cost Analysis, provide a very powerful tool for scenario planning. It is interesting to note that such applications give rise to many interesting planning problems. A number of Banks have found that having carried out their Corporate Planning approach on large planning Models, a small number of important alternative strategies emerge. These give rise to the classical O.R. approach of having small inter-disciplinary teams investigating each area of these potential strategies. I am pleased to report in one major European Bank such teams are led by Managers who have worked for a number of years in the Bank's O.R. department.

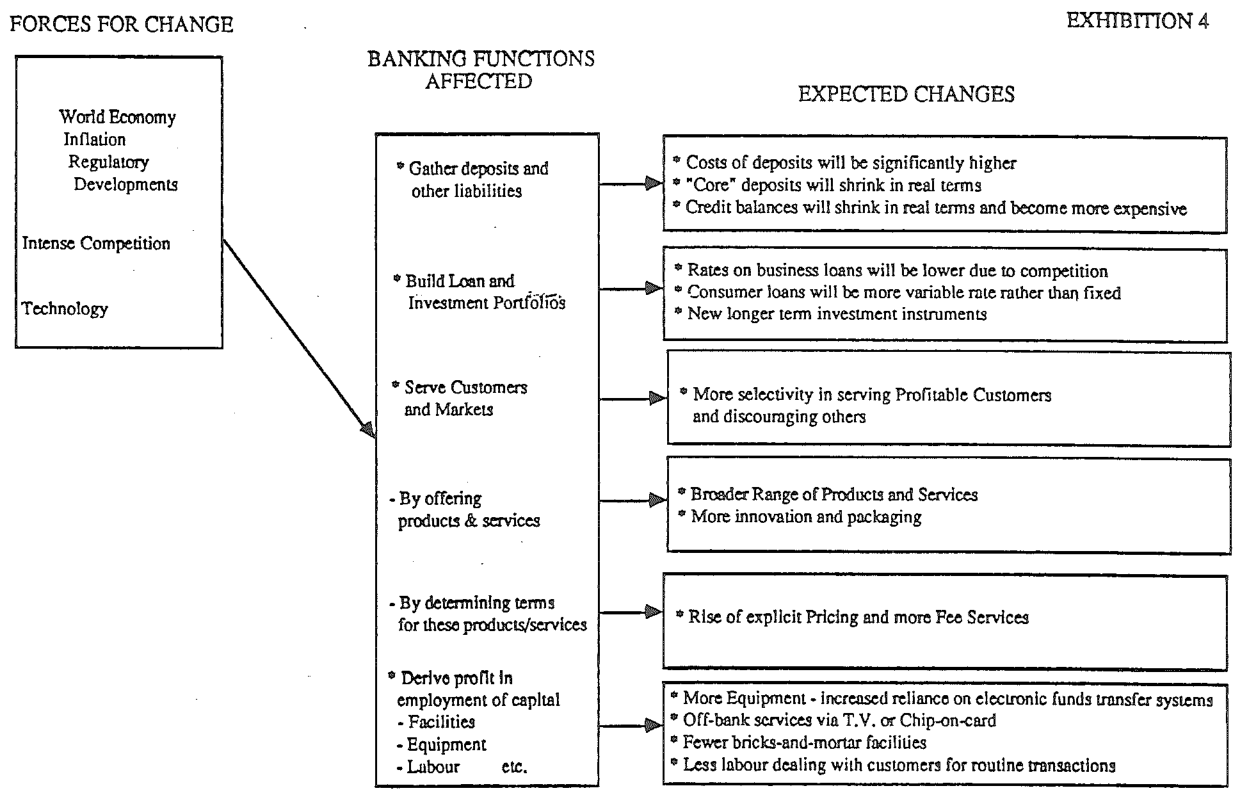
7. Current Active Work Areas for O.R. in Banking

7.1 Branch Network Development

European Banks are now faced with major strategic decisions with regard to the type of delivery systems they provide for their services. Traditionally, the Branch network with full service branch outlets was the way in which banks delivered their services. However, with the advent of automation and self-service banking a lot of reasons for the full service branch network are now coming into question. It is vitally important for Bank Management to fully understand the options for new delivery systems and to begin to make important decisions on the shape of their future networks. This is an area in which Operations Research is playing a major role. This is truly an experimental area and many banks are attacking it using the classical O.R. approach. For example, via development of simulation models of a branch bank to determine the impact of new technologies on delivery systems. The next stage is to actually implement some of these alternative strategies in experimental branches. And finally, to use O.R. to assess the results of these

EXPECTED CHANGES IN EUROPEAN BANKING

EXHIBITION 4



experiments. A very new application in this area is the use of Micro-computers in the simulation. New techniques now allow what is called slow speed simulation with excellent graphic displays. I have seen models of a branch banking office in which you can see customer's arrivals, queuing for service and leaving the branch displayed on a V.D.U. Such models can be run at different speeds and can accumulate for a given period, all the queuing statistics pertinent to the branches operation.

The other major area of interest at the moment is in the Marketing Area. Due to intense competition, not alone from other bank but from other Financial Institutions entering the financial services market, it is most important for the banks to understand their market, to identify market segments and to develop marketing strategies to meet these segments needs. The availability of large customer data bases is now becoming a reality and this gives important applications for advanced statistical techniques-such as cluster analysis in the determination of market segments. Here again, work done by Operations Research in costing of products/services can be combined with Market Research techniques to produce profitability by market segment.

8. Future Development

Exhibit 4 shows the expected changes in European Banking over the next decade. European Banks are being assailed by competition from other Financial Institutions and Foreign Banks. Banking costs and new technologies, possible major regulatory changes will all further change the look and practice of banking in the next decade. The foremost problem will be, the difficulty of maintaining adequate levels of capital and profitability. Nevertheless, in spite of uncertainty the 1980's will hold opportunities as well as threats for the Banks.

It is essential for the Banks to fully utilise the skills of Operations Research in responding to the challenge ahead. In particular, I feel the areas of Marketing and Branch Delivery Systems will offer major opportunities for the O.R. worker to make a contribution. For example, at the moment in France there are a number of major experiments with the use of the 'Chip Card'. This is a new device which looks like a Credit Card but contains a Microchip which can store information and be modified by transactions. A costumer can have money transferred to this 'Chip Card' and then use this Card to pay shops and other businesses by having the Card pass through a simple device in the shop. This is basically an electronic wallet. It means that an off-line system can be operated for shops and retailers. If such a system proved to be satisfactory, clearly the implications for a branch banking network would be very important. This is an application where Operations Research can be used to model alternative scenarios and support Management in a very difficult and major decision area.

In summary, I believe that O.R. has made a considerable contribution to the development of banking both in Europe and in the United States. It is currently very active and with the competitive challenge ahead for banks it will be vital to utilise the skills and professionalism available trough Operations Research.

Investigação Operacional, Vol.7 nº1 (Junho/1987) 11 - 16

APPLICATIONS OF OR IN PUBLIC SERVICES PLANNING

Tom Bowen
Department of Health and Social Security, London, UK

I was very pleased to be invited to address this conference on the subject of the contribution that Operacional Research can make to the planning of public health services. My talk is largely concerned with approaches that have been developed in the United Kingdom (UK). Nevertheless, the methodologies that we have developed are sufficiently flexible to allow application in other countries, and I would hope that in Portugal itself the development of a National Health Service would be greatly assisted by the application of OR methods.

In the UK, health planning studies are often described as either operational or strategic, and OR is used for both. Operacional studies are those concerned with immediate or short-term problems and are generally carried out from the "bottom up". The problems encounted may be concerned with administrative management (eg stock control, vehicle replacement), for which the OR approach will not differ greatly from that adopted in non-health organisations; or the problems may be specifically concerned with health service delivery eg how long does a particular type of patient need to stay in hospital? How many staff are needed to provide adequate cover in a casuality ward? How should outpatient appoitments be scheduled?

Today I would like to concentrate on strategic planning, which is concerned with resolving long-term problems and generally requires a "top-down" approach. Here the Operational Research is faced with a major and often highly interdependent questions: how many patients need any particular service? How many hospital beds are needed? What will be the service implications of increasing numbers of elderly people? How many staff do we need to train now to provide for a desired future service level?

Before discussing how some of these questions are tackled in the UK I will give a little background. Total expenditure per annum on health and social services in the UK is currently around $30 billion, or about $600 per capita. Approximately 75% of this is incurred by the National Health Service (NHS), but other agencies are also envolved: a further 15% is spent by local authorities on personal social services (residential care, social workers, etc) and there is also a small (but growing) amount of private and voluntary provision.

What do these agencies do? People commonly think of health services in terms of doctors providing medical treatments. I would however tend to emphasize the role that health organisations have in the prevention of ill-health, for example by improved ante-natal care, immunisation programmes and educational campaigns to improve diets or reduce cigarette smoking. If preventive measures are unsuccessful then treatment becomes necessary, and here we have to consider a vast range of activity, from open-heart surgery to providing aspirin for the common cold. And when treatments fail, or to when none are available, patients may need long-term care, to enable them to continue to live lives that are as normal as possible.

The OR Approach

We could consider prevention, treatment and care to be the three objectives of any health system. In any OR study we would of course wish to specify the objectives more precisely than this. There are then perhaps three further steps which encapsulate the OR approach.

- specify a set of options from which a choice may be made;
- develop a means to categorise or measure the options in terms of the stated objectives;
- develop logical procedures (models) to evaluate the desirability of each option and hence planning decisions.

As an example of these aspects of the OR approach I shall consider the long-term care of the elderly. In other words, how can OR help in planning better care for dependent elderly people? Table 1 shows some of the relevant components for an OR study: note that each criterion, eg "higher standard of care", needs to be more thoroughly specified as the study progresses.

Table 1 also shows some of the problems commonly encountered in health OR. Perhaps most important is the phenomenon of the "multi-organisation": in health systems not only may there be several agencies involved in providing services, but even within a single agency there may be conflicts between professional groups (doctors, nurses, social workers, etc) over the objectives of the service. Patients and their relatives may also have their independent aims. Finally, and frequently forgotten, are the members of the population who are healthy, and wish the health system to act in such a way as to preserve their good health. The immediate consequence of the "multi-organisation" is the need for the OR analyst to stimulate negotiation between the parties concerned, to try, as a first step, to reach a consensus over the options that might be acceptable to move towards health objectives.

The other major dilema for the OR analyst concerns measurement. There may be problems in gathering data of sufficient quality to support a quantified or systematic analysis. If data are available, then there are often problems in expressing the output of the system: there is not, and is not likely to be in the near future, any acceptable means of measuring "health status" on a routine basis for the general population, which would be essential if we wished to quantify the extent to which health services actually promote good health. In practice, OR studies tend to talk of intermediate outputs, for example the number of patients treated in a year in a given speciality. Much OR work revolves around the determination of the inputs (hospital beds, doctors, etc) that are required to achieve such intermediate outputs.

Balance of Care

The presence of all these problems make health OR particularly interesting, and the Balance of Care approach highlights many of the problems discussed above. This approach has developed over some twelve years in the UK and after application in several health and local authorities has been found of great use in planning for the long-term care needs of the elderly.

However, the essential philosophy of the method has been carried across in studies of other patient groups.

The structure of the problem is shown in figure 1. The patient group, the elderly, is sub-divided into categories with varying needs for health services. These can be met by a combination of services, which in turn could be supplied by more than one agency*. For a given category there may be several options, called modes of care, for meeting the individual's needs.

STRUCTURE OF THE PROBLEM

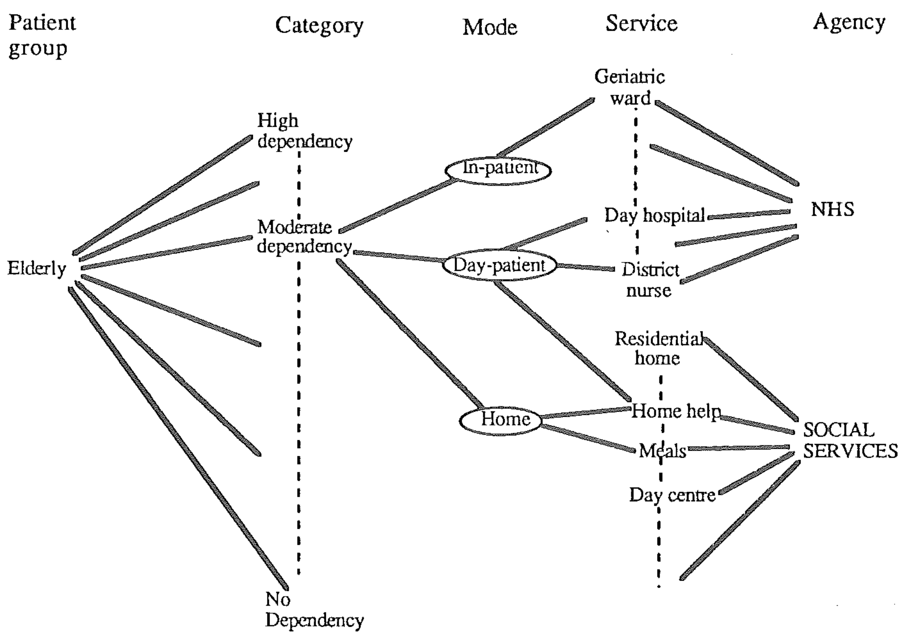


Figure 1

TABLE 1 - STEPS IN THE OR APPROACH

	Example	Common Problems
Objectives	Better care of the elderly. Higher standards of care. Services for all in need. Economy	Multi-organisations
Options	Care in the community Care in the institutions	No consensus
Categorise	Types of patient Current provision	Information not available, too subjective or unquantifiable
Evaluate	Future needs of aging population Cost - effective service mixes	No apropriate measures of output

* In the UK we are normally concerned with the two biggest agencies, the NHS and the Local Authority Social Services, and applications of Balance of Care are particulary concerned to assist the relevant authorities in the joint planning of services.

The first major stage in a Balance of Care exercise is to specify what characteristics in elderly patients lead to different service needs (ie define the categories) and what modes of care could satisfy those needs. This task is undertaken by a group of professional advisors, each of whom will be responsible for one aspect of service delivery. Thus, for the elderly, the group might consist of a geriatrician, a psychiatrist with special interest in the elderly, a district nurse, a social work team leader, a residential homes manager and a home help organiser.

A simplified example of the type of category that such a group will define, and the modes of care thought appropriate is given in figure 2. Four possible modes of care are proposed here: hospital care in a long-stay geriatric bed, care in a residential home, and care provided in a patient's own home, with or without regular attendance at a day hospital. Figure 2 also shows the advisory group's order of preferences for each mode of care: on professional grounds the group would prefer to care for the patient in his or her own home, but with the support provided by regular day hospital attendance. However, the overall cost of this option is quite high, so would not necessarily be the cost-effective choice. In figure 2 it is also worth noting that the cost to the agencies providing the services is often less (sometimes substantially less) than the overall cost: this is because the individual, often via the social security system, may have to meet part of the total cost. Thus under each of the three criteria shown here (professional preference, total cost, cost to agencies) a different mode of care will be chosen.

In the UK there is no routinely collected data giving the type of information needed to quantify this structure. A major feature of Balance of Care studies is therefore to organise surveys to find out how many of each category of patient exists in the district being studied, and to determine what services are being provided and in what quantity. Normally the data are collected by questionnaires, which are filled in by the staff providing the services. Sometimes samples are used to reduce data collection costs.

FIGURE 2 - EXAMPLE OF A PATIENT CATEGORY

Characteristics				
MENTAL	No dementia			
PHYSICAL	Severe physical disability			
CONTINENCE	Occasional incontinence			
SOCIAL	Support from family			
Modes of Care				
Long-stay geriatric bed	X			
Residential home		X		
Day hospital			X	
District nurse			X	X
Home help			X	X
Preference	2	4	1	3
Cost per week	$350	$280	$300	$200
Cost to agencies per week	$350	$ 80	$200	$100

There are some immediate benefits from the surveys. Firstly the information can be fed back to the advisory group, to enable them to reconsider and refine their earlier judgements of what constitutes an acceptable or desirable standards of care. Secondly, this feedbeck may indicate major allocative inefficiency (eg a particular service may be given in large quantity to patients who only need a little, while those most in need are neglected): this is an operational planning problem, which needs to be considered separately, but in conjunction with any existing rules govering the eligibility of patients for particular services (eg criteria for admission to a residencial home).

The major use of the data however lies in the modelling approach to the development of strategic plans. This is shown in outline in figure 3. Current demand and use of services is derived from the surveys and future demand estimated on the basis of projected population changes. Cost effective plans for future services are then developed from a sequence of calculations of either an optimising or simulating type. Optimising is used to determine and qualify the "best" services to provide to meet policy objectives in

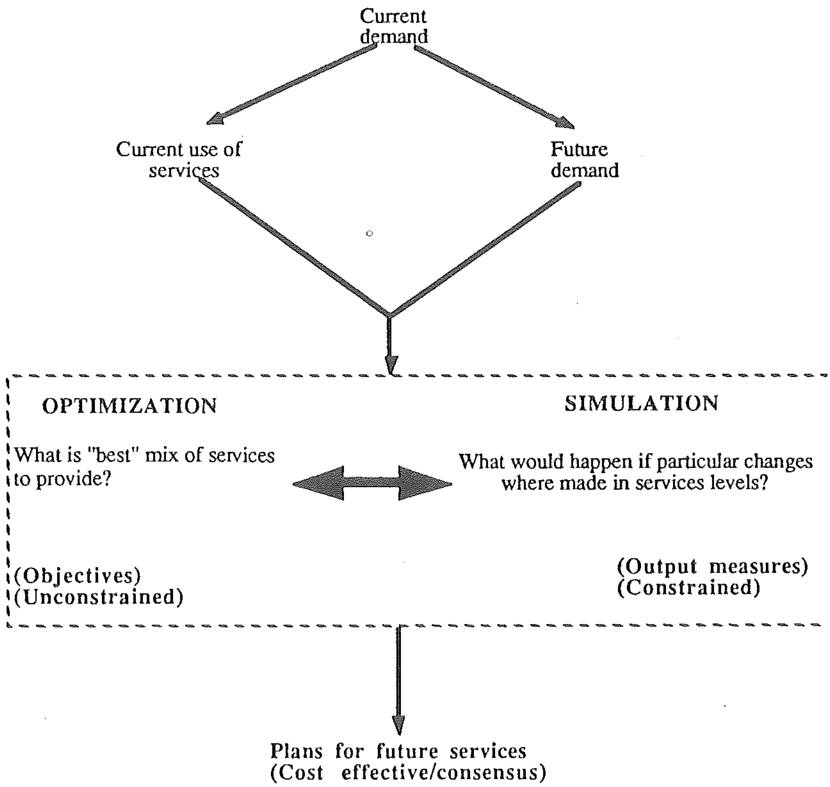


FIGURE 3 - OUTLINE OF THE MODELLING APPROACH

an unconstrained world. Since there are many objectives (high standards, economy, etc), "best" may be interpreted in different ways and thus many solutions are possible. These solutions are intended to help planners develop planning "scenarios" which can be tested using simulation analyses: here we are concerned to determine <u>what</u> would happen <u>if</u> particular changes were made in service levels. Of much interest here is the impact that different patterns of service have on the numbers receiving a service and the standard of that service (relative to the levels prescribed by the advisory group): these are out measures of intermediate output.

To be successful this aproach needs significant involvement by both planners and professionals. In practice we would expect them to make all the major assumptions for each analysis. By doing so, the key ingredients of the strategic plan gradually fall into plaĉe and degree of consensus is reached between the various interested parties.

<u>Discussion</u>

What are the lessons for OR projects in the planning of public health services? Firstly that top-down approaches, such as the one I have briefly described here, can have a major impact on the strategic direction of an organisation. "Traditional" OR studies, which I have described under the label "operational planning" have been equally successful, but will usually be carried out for lower levels of management and have fewer resource implications.

A second lesson that we have learnt concerns the importance of data collection. Coherent planning cannot take place without relevant information, and sometimes elementary statistical analyses will contribute as much to the planning process as the more sophisticated modelling that is a feature of the Balance of Care approach. That said, it is important to recognise that the data collection for Balance of Care studies takes place within a structured framework - so we can be sure that the data collected will be useful!

My final point concerns the environment in which all this work is carried out. Good Operational Research is never carried out in a "back-room" and, in the health sector, will require the active involvement of both planners and professional staff. Since however individual objectives are likely to conflict, the OR analyst will find himself involved in a process of negotiation between the participants, with particular analyses often being supported only by those likely to gain from their implementation (eg a district nurse organiser is likely to support an analysis indicating that more district nurses should be empolyed!) Indeed, by making the various issues more explicit, the OR analyst may find that decision making has been made <u>more</u> difficult - but in the end better decisions will be made!

References on "Balance of Care"

1. KLEMPERER P. D. & McCLENAHAN J .W. (1981). Joint strategic planning between health and local authorities. Omega, Vol. 9, Nº 5, pp. 481-491.
2. BORLEY R., TAYLOR S. & WEST C.R. (1981). A user's view of a new approach to joint strategic planning. Omega, Vol. 9, Nº 5, pp. 493-499.
3. CANVIN R., HANSON J., LYONS J. & RUSSELL J. C. (1978). Balance of Care in Devon: joint strategic planning of health and social services at AHA and County level. Health & Social Services Journal, Vol. 88, pp. C17-C20.
4. NICHOLLS I. G. (1981). Strategic planning of health and personal social services in Dudley - experiences with the Balance of Care model. (Proceedings of the Second International Conference on Systems Science in Health Care, Montreal, 1980).
5. GIBBS R. J. (1978). The use of a strategic planning model for health and personal social services. Journal of the Operational Research Society, Vol. 29, Nº 9, pp. 875-883.
6. McDONALD A. G. , CUDDEFORD G. C. & BEALE E. M. L. (1974). Mathematical Models of the Balance of Care. British Medical Bulletin, Vol. 30, Nº 3, pp. 262-270.

Investigação Operacional, Vol.7 nº1 (Junho/1987) 17-25

SYMMETRIC CODES FOR REVISED SIMPLEX TECHNIQUES

Heiner Muller - Merbach
President of IFORS
Universitat Kaiserslautern, Operations Research

Summary: A frame for revised simplex techniques will be presented which is based on a row column symmetry. The frame includes the LU representation of the core inverse. The advantage of symmetry is that the computer codes consists of only a small number of subroutines each of which is applied twice per iteration. This frame is particulary advisable for micro computers.

1. Introduction: Advantages of Symmetry

It is the magic number of 37 years ago that the simplex technique was created by George B. Dantzig. Many variations of this technique followed since. Among them is the "revised" simplex technique in different versions.

Most of the presentations of the (standard and revised) simplex technique in the textbooks, the monographs and in the OR journals are column oriented. This applies to the EDP standard codes of linear programming (LP) as well.

Only a minority of authors plea for symmetric versions of the simplex technique. These versions are row and column oriented to the same extent.

Symmetric versions of the simplex technique do have some advantages over column oriented techniques - in addition to the aesthetic beauty:

(i) Many subroutines can be used twice, for rows as well as for columns, and the number of subroutines can be reduced accordingly.

(ii) An identity matrix which is a characteristic property of the column oriented technique becomes obsolete.

(iii) Primal and Dual pivot selection rules can be interchanged without any adaptation to the column orientation.

(iv) Bounding techniques (lower, upper, general, variable bounds) can be dualised and implemented without adaptation to the column orientation.

(v) While the column oriented versions allow only late insertions of additional columns (while additional rows are quite difficul to insert), the symmetric techniques allow for late insertions of rows and columns in the same way. This is important with respect to cutting planes and obliquely angled branch and bound (see Muller - Merbach 1983a).

It is surprising that the symmetric versions of the simplex techniques have not found the support of the majority among the mathematical programming experts.

In this paper, a brief survey of symmetric versions of the simplex technique shall be presented. The symmetric standard technique will be presented in section 2. This will be extended in section 3 in which block pivoting is introduced. Based upon the block pivoting formulae, a frame of symmetric revised simplex techniques will be given in section 4. Its kernel is the core inverse, replacing the tradicional basis inverse. LU representations of the core inverse follow in section 5. Manipulations with L and U and their inverses are presented in section 6, in which the notation of a matrix division operator is introduced. The paper terminates with conclusions in section 7.

2. The Symmetric Standard Simplex Technique

The following notation for linear programming (LP) models shall be used here:

$$\begin{aligned} &\text{Maximise} \quad y_o \\ &y_o + c'x = d \\ &y \;\; + Ax = b \qquad\qquad (1) \\ &x, y \geq 0 \end{aligned}$$

In this notation, c' and x are n-vectors, y and b are m-vectors, and A is an mxn-matrix.

For some purpose, it is convenient to rewrite system (1) in the following way, using the equivalences $a_{oj} = c_j$, $a_{io} = b_i$ and $a_{oo} = d$:

$$\text{Maximise} \quad y_o$$

$$y_i + \sum_{j=1}^{j=n} a_{ij} \quad x_j = a_{io} \qquad \text{for } i = 0, 1, \ldots, m \qquad (2)$$

$$x_j \geq 0 \qquad \text{for } j = 1, 2, \ldots, n$$

$$y_i \geq 0 \qquad \text{for } i = 1, 2, \ldots, m$$

The x_j are initial **"non-basic" variables**, while the y_i are initial **"basic" variables**. During each iteration of the simplex technique, one basic variable will be exchanged against one non-basic variable.

Consequently, the coefficients a_{ij} change with each iteration. In the following formulae, an upper index p shall indicate the values of the single data after the p-th iteration. Furthermore, the current basic variables will be represented by BV_i, the current non-basic variables by NBV_j. Thus, the equation of (2) can be replaced by the more general formulation:

$$BV_i^{(p)} + \sum_{j=1}^{j=n} a_{ij}^{(p)} NBV_j^{(p)} = a_{io}^{(p)} \qquad \text{for } i = 0, 1, \ldots, m \qquad (3)$$

In each single simplex iteration, a pivot column s and a pivot row r have to be selected. After their selection, the system (3) will be updated according to the following rules:

a) Pivot element: $a_{rs}^{(p)} := 1/a_{rs}^{(p-1)}$ (4a)

b) Pivot row: $a_{rj}^{(p)} := a_{rs}^{(p)} a_{rj}^{(p-1)}$ for $j \neq s$ (4b)

c) Pivot column: $a_{is}^{(p)} := -a_{is}^{(p)} a_{rs}^{(p)}$ for $i \neq r$ (4c)

d) Remaining matrix: $a_{ij}^{(p)} := a_{ij}^{(p-1)} - a_{is}^{(p-1)} a_{rs}^{(p)} a_{rj}^{(p-1)}$ for $i \neq r, j \neq s$ (4d)

Rule (4d) can be simplified by application of (4b) or (4c).

e) Exchange of variables: $BV_r^{(p)} := NBV_s^{(p-1)}$

$NBV_s^{(p)} := BV_r^{(p-1)}$ (4e)

These formulae are used by authors such as Muller-Merbach (1965, 1967, 1973), Stiefel (1963), Vajda (1961, 1981), Wiezorke (1961), Zoutendijk (1976) and others. The symmetry lies in the formulae (4b) and (4c). Most authors, however, extend the matrix A of (1) by an identity matrix carrying the coefficients of the basic variables explicitly. They do not need the rules (4a) and (4c), because the operation of (4a) would be covered by (4b) and the operations of (4c) by (4d).

For those readers who are not as yet familiar with symmetric standard simplex technique, a numerical example will be presented, taken from Muller-Merbach (1973). The model - written in tableau form - reads:

	X_1	X_2	
y_o	-300	-500	-36000
y_1	1	2	170
y_2	1	1	150
y_3		3	180

The first iteration would lead to the exchange of X_2 against y_3:

	X_1	y_3	
y_o	-300	500/3	-6000
y_1	1	-2/3	50
y_2	1	-1/3	90
X_2		1/3	60

The second iteration would lead to an exchange of X_1 against y_1:

	y_1	y_3	
y_o	300	-100/3	9000
x_1	1	-2/3	50
y_2	-1	1/3	40
X_2		1/3	60

The third and last iteration would bring y_3 back into the basis which is left by y_2:

	y_1	y_2	
y_o	200	100	13000
X_1	-1	2	130
y_3	-3	3	120
X_2	1	-1	20

3. Symmetric Block Pivoting

After any number of iterations, a certain number of initial basic variables are non-basic, and the same number of initial non-basic variables are basic. Based upon these variables which have entered or left the basis, the single element pivoting formulae (4a-e) can be replaced by the following block pivoting formulae (7a-e) which comprise all the operations of the intermediate single iterations.

The following block indices are to be used for block pivoting:

R : Rows with an initial non-basic variable as current basic variable.

I : Rows with an initial basic variable as current basic variable (including the objective function).

S : Comuns with an initial basic variable as current non-basic variable.

J : Columns with an initial non-basic variable as current non-basic variables (including the right hand side).

With these block indices, the initial LP system (2) can - with respect to any state of iteration - be partitioned in the following way:

	NBV_S	NBV_J
BV_I	A_{IS}	A_{IJ}
BV_R	A_{RS}	A_{RJ}

(5)

Since the basic and non-basic variables change with any iteration, the partitioning of (5) would have to be adapted after any iteration. Through the formulae (4a - e) this initial system would after any iteration p - turn into:

$$\begin{array}{c|cc|} & NBV_S^{(p)} & NBV_J^{(p)} \\ \hline BV_I^{(p)} & A_{IS}^{(p)} & A_{IJ}^{(p)} \\ BV_R^{(p)} & A_{RS}^{(p)} & A_{RJ}^{(p)} \end{array} \qquad (6)$$

The relation between (6) and (5) can be expressed by the following block pivoting formulae, which are generalisations of (4a - e):

a) Pivot block: $A_{RS}^{(p)} := A_{RS}^{-1}$ (7a)

b) Block pivot row: $A_{RJ}^{(p)} := A_{RS}^{(p)} A_{RJ}$ (7b)

This includes part R of the right hand side.

c) Pivot block column: $A_{IS}^{(p)} := -A_{IS} A_{RS}^{(p)}$ (7c)

This includes part S of the objective function.

d) Remaining matrix: $A_{IJ}^{(p)} := A_{IJ} - A_{IS} A_{RS}^{(p)} A_{RJ}$ (7d1)

or $A_{IJ} - A_{IS} A_{RJ}^{(p)}$ (7d2)

or $A_{IJ} + A_{IS}^{(p)} A_{RJ}$ (7d3)

This includes part I of the right hand side and part J of the objective function.

e) Exchange of variables $BV_R^{(p)} := NBV_S$ (7e)

$NBV_S^{(p)} := BV_R$

Such block pivoting formulae are well-known from the literature, see Dantzig (1963a, p. 201), Land and Powell (1973, pp. 15ff.), Muller-Merbach (1966, 1967, 1982, 1983b), Orchard-Hays (1968, pp. 27ff.) etc. In most cases, however, the block pivoting formulae are slightly different from (7a - e) because they are referred to a column oriented formulation of the LP model which includes an additional identity matrix, see the discussion of formulae (4a - e) in section 2.

Through the block pivoting formulae (7a - e) the property of symmetry is maintained, see (7b) and (7c) as well as (7e). This is essential.

The key role in (7a - d) plays the submatrix ARS and its inverse. They shall be called core matrix and core inverse, respectively. This core inverse is in a slight contrast to the basic inverse, a standard term in the LP literature and the key submatrix in most versions of the revised simplex technique. The basis inverse consists of the core inverse (7a) plus the pivot block column (7c) plus an identity matrix (with the same row number as the block pivot column). The basis inverse corresponds with the column oriented organisation of the simplex technique while the core inverse corresponds with the symmetric versions.

Core inverse: $A_{RS}^{(p)}$ Basis inverse: $\begin{array}{cc} A_{RS}^{(p)} & 0 \\ A_{IS}^{(p)} & I \end{array}$

4. Symmetric Revised Simplex Techniques, Based on the Core Inverse

As suggested by the author since 1966 (see Muller-Merbach 1966, 1967, 1982, 1983b) revised simplex techniques can efficiently be organised in a symmetric way. The symmetry appears in the equivalent treatment

- of the right hand side and the objective function,
- of the pivot column and the pivot row,
- of the basic variables and the non-basic variables.

The basic idea that led to the revised simplex techniques was that not all the matrix A, see system (2), has to be kept updated throughout all the iterations. Those parts of A which are not kept updated have to be computed "from scratch" when required. The main difference between the single versions of the revised simplex technique lie in the question which parts of A are kept updated.

In any case, the core inverse has to be kept updated (either explicity or in LU form, see section 5). In addition, it is suggested here to keep the full right hand side and the full objective function updated. It is further suggested to keep the pivot block row A_{RJ} (except for part R of the right hand side), the pivot block column A_{IS}(except for part S of the objective function) and the remaining matrix A_{IJ} (except for part I of the right hand side and for part J of the objective function) untouched. Their actual values have to be computed from scratch when required; since the rules (7b - d) can be applied to the full matrix as well as to any single row or column or element, the computation from scratch can be carried out in a selective manner.

Thus, a general frame for the symmetric revised simplex technique can be stated:

- Step 1: Select pivot column s according to the objective function.
- Step 2: Compute pivot column:

  If $s \in J$,
  then compute from scratch part R of pivot column acc. to (7b) and part I acc. to (7d2).

  If $s \in J$,
  then copy part R of the pivot column from the core inverse (if in explicit form: update if in LU form), and compute from scratch part I of the pivot column acc. to (7c).
- Step 3: Select pivot row r according to the right hand side and the updated pivot column.
- Step 4: Compute pivot row:

  If $r \in I$,
  then computer from scratch part S of the pivot row acc. to (7c) and part J acc. to (7d3).

  If $r \in R$,
  Then copy part S of the pivot row from the core inverse (if in explicit form: update if in LU form), and compute from scratch part J of the pivot row acc. to (7b).
- Step 5: Update the system:
  - a) Update the core inverse acc. to (4a - d) using part R of the pivot column and part S of the pivot row and the pivot element.
  - b) Update right hand side (part R and I) acc. to (4b and d).
  - c) Update objective function (part S and J) acc. to (4c and d).
  - d) Exchange variables acc. to (4e).

  Return to step 1.

Exits (optimal solution, no feasible solution, infinite solution) would have to be inserted in step 1 and 3.

This frame of the revised simplex technique has several symmetric properties, such as:

- Step 1 and step 3 are almost equivalent and can be controlled by the same subroutine.
- Step 2 and step 4 are fully symmetric to each other and can be carried out by the

same subroutine.

- Part a) of step 5 is symmetric in itself, due to the symmetric between (4b) and (4c).
- Part b) and part c) of step 5 are symmetric to each other and can be carried out by the same subroutine.
- Part d) of step 5 is symmetric in itself, due to the symmetry of (4e).

The suggested frame of the symmetric revised simplex technique contains at least one suggestion which is quite in opposition to traditions of the revised simplex technique. It refers to part c of step 5, namely to keep the objective function updated. It is common use in revised simplex techniques to compute the objective function from scratch in every iteration, i.e. (i) to compute from scratch the simplex multipliers (part S of the objective function) according to (7c) and then (ii) to compute from scratch part J of the objective function according to (7d3). If it is done in the way, one does not need to compute part J of the pivot row (step 4). Thus, why is it suggested here to keep the objective function updated throughout all the iterations? The reaction is simple. To keep the objective function updated, one has to compute from scratch the full pivot row. Part R of the pivot row would be required anyway, namely for part a) of step 5. This part can be used to keep at least part S of the objective function (simplex multipliers) updated. Thus, the alternative remains with respect to part J of the objective function. In case of density 1 for A (i.e. no zero elements) there would be no siginicant computational difference because in any way, part J of a row (objective function or pivot row) has to be computed from scratch. However, usually the single rows are much less dense in the average than the objective function. Therefore, it takes - in practice - less operations to compute from scratch part J of the pivot row than part of the objective function.

The only open question lies in the representation of the core inverse which will be considered now.

5. The LU Representation of the Core Inverse

It is well-known since Gauss (1777-1855) that a non-singular matrix A can be represented by the multiplication of two triangular matrices, a lower triangular matrix L, and a upper triangular matrix U, namely: $A = LU$. This can be applied to the inverse as well: $A^{-1} = U^{-1} L^{-1}$. It has been suggested by Markowitz (1957), Dantzig (1963b), Bartels and Golub (1969), Forrest and Tomlin (1972), Powell (1975) and others to apply the LU representation of the inverse to the revised simplex technique as well. In the early cases, this representation was applied to the basis inverse, in the more recent cases to the core inverse only. But the column oriented concept of the basis inverse was being maintained in most of these publications.

In this paper it is proposed to combine the LU representation with the symmetric version of the revised simplex technique. This only requires to replace the core inverse by

$$A_{RS}^{(p)} = U_{RS}^{-1} L_{RS}^{-1} \tag{8}$$

A symmetric revised simplex technique working literally (even if a different notation was used) with (8) was published by Muller-Merbach (1965) and (also in a different notation but with the same content) some years later by Haverly (1972). This representation of thr core inverse is extremely simple, computationally:

- The upper triangular inverse would grow by one column (and zero row) in each iteration. This new column is identical with part R of the pivot column (plus pivot element). Except for the additional column and row, the previous upper triangular matrix would in principle remain unchanged.
- The lower triangular inverse would grow by one row (and zero column) in each iteration. This row is identical with part S of the pivot row (plus a 1 in the diagonal). Except for the additional column and row, the previous lower triangular matrix would in principle remain unchanged.

Particular attention has to be paid to the case that the pivot element does not come from the submatrix A_{IJ}, see system (5 and 6). If the pivot element comes from A_{IJ}, the core inverse would grow by one row and one column, as do L_{RS} and U_{RS} and their inverses.

However, if the pivot element comes from the core matrix A_{RS} itself, the core matrix and its inverse would shrink by one row and one column. If the pivot element comes from A_{RJ} or A_{IS}, the core inverse would keep its size. In these cases, there exists two possibilities:

a) Extension: The inverse of L_{RS} and U_{RS} could be allowed to grow by one row and column in each iteration, idependently of the actual size of the corresponding (virtual) core inverse. This way was taken by Muller-Merbach (1965) and Haverly (1972).

b) Adaptation: The shrinking or stagnation of the core inverse can be represented by a corresponding shrinking or stagnation of the triangular inverses. This, however, would require an adaptation of the triangular matrices themselves which would be kept unchanged otherwise.

The replacement of the core inverse according to (8), i.e. by the triangular inverses was not the suggestion of the protagonists of the LU representation of the basis inverse or core inverse, respectively. Instead, they suggested to work immediately with L and U, and not with their inverses. This has the advantages that L and U are usually much less dense than their inverses. For their convenient representation in algorithms, the matrix division operator is advantageous. It will be introduced in the following section.

6. The Matrix Division Operator

It is well-known that a division by a matrix is not possible in principle. An exception are triangular matrices even if it is not common use to speak of a matrix division. Rather, "elimination" is the common term. However, if "division" is defined by dividing a dividend by a divisor without changing the divisor, then elimination can be understood as division, as far as triangular matrices are concerned. For them, two versions of a matrix division operator shall be introduced:

: represents the division of a matrix or a vector by a triangular matrix to the right.

: represents the division of a matrix or a vector by a triangular matrix to the left.

Both versions of this operator shall be explained now. Consider the system

$$Ax = b \tag{9}$$

with x and b as p-vectors and A as pxp-matrix. System (9) can be rewritten as

$$x = A^{-1}b = U^{-1} L^{-1} b = U: (L : b) \tag{10}$$

Here, the multiplication of the column vector b by the triangular inverses is replaced by its division by the triangular matrices to the left.

Consider the system

$$yA = c \tag{11}$$

with y and c as row vectors. This system can be rewritten as

$$y = cA^{-1} = cU^{-1} L^{-1} = (c : U) : L \tag{12}$$

Here, the multiplication of the row vector c by the triangular inverses is replaced by its division by the triangular matrices to the right.

These matrix divisions of type (10) and (12) can be implemented everywhere in (7a - d) in replacement of the core inverse.

Which operations are behind the matrix division operator? It is nothing else than elimination. Four cases have to be considered:

(i) Division of a column vector by a lower triangular matrix (to the left).
(ii) Division of a column vector by an upper triangular matrix (to the left).
(iii) Division of a row vector by an upper triangular matrix (to the right).
(iv) Division for a row vector by a lower triangular matrix (to the right).

Only for case (i) the single operations shall be made explicit. The other cases are similar. Consider the system

$$Lz = b \tag{13}$$

in which L is lower triangular matrix with $l_{ij} = 0$ for $j > i$. System (13) represents the following single operations:

$$\sum_{j=1}^{j=i} l_{ij} z_j = b_i \qquad \text{for} \qquad i = 1, 2, \ldots, p \tag{14}$$

The system can be solved for z_i according to

$$z_i = (b_i - \sum_{j=1}^{j=i-1} l_{ij} z_j) / l_{ii} \quad \text{for} \quad i = 1, 2, \ldots, p \tag{15}$$

This operation (carried out in the sequence i = 1, 2,..., p) represents precisely the elimination procedure for which the matrix division operator is suggested here:

$$z = L : b \tag{16}$$

Thus, the operations are common use, the operator not. The operator, however, makes the handling of L and U highly convenient.

Almost the same procedures apply to the three other cases. There is a two-dimensional symmetry between the four cases. All of them would be handled by the same subroutine controlled by two parameters.

With reference to the symmetric revised simplex technique drafted in the frame of section 4, the representation of the core inverse by a pair of triangular matrices (L and U or their inverses) affects in particular step 2 and 4 as well as part a) of step 5. In step 2, the division of type (10) would replace the multiplication by the core inverse in (7b) if $s \in J$; if $s \in S$, an updating of part R of the pivot column would be necessary, since the core inverse is not explicitly available. In step 4, the division of type (12) would replace the multiplication by the core inverse in (7c) if $r \in I$; if $r \in R$, anupdating of part S of the pivot row would be necessary, since the core inverse is not explicitly available. In part a) of step 5, not the fully updated part R ρf the pivot column and part S of the pivot row would be needed. Instead, part R of the pivot column (in its initial state) would have to be divided by U_{RS} (to the left), while part S of the pivot row (in its initial state) would have to be divided by L_{RS} (to the right). These new vectors would be added as new column or row of U_{RS} and L_{RS}, respectively. Should the core inverse shrink or stagnate in size, there exists the choice between extension and adaptation, as outlined in section 5.

The LU representation of the core inverse has computational advantages. They depend in particular upon density considerations. Were the matrix A (as well as the corresponding triangular matrices) filled with non-zero elements (density = 1), procedures of type (10) and (12) would in any case require p^2 single operations. This is different however if A is sparse. If for instance a 100x1000-matrix A has an initial density of, say, 0.03 (randomly distributed), L and U would have an average density of 0.25, their inverses a density of 0.72, and the explicit inverse A^{-1} a density of 1. This follows from some theoretical experiments. Procedures of type (15) would in this case reduce the number of operations by three quarters. In addition, much less data have to be stored.

This is the reason why modern standard EDP codes of linear programming use the LU representation of the inverse and why they work with the elimination technique. However, they are still column oriented and do not take advantage of the properties of symmetry. In particular, they do not keep the objective function updated; instead, they compute from scratch the simplex multipliers (part S of the objective function) and use them for computing from scratch the remaining part J of the objective function.

7. Conclusion

This paper was a plea for symmetric codes of the revised simplex technique. Symmetry of the codes is nothing else than immediate application of duality theory.

The algorithmic symmetry has the advantage that the revised simplex technique can be reduced to a very small number of subroutines, each of which can be used twice or even fourth. Therefore, the EDP codes would become small which is an interesting property for micro computers.

In addition, the symmetric structures allow for the implementation of dual bounding techniques in the same way as primal bounding techniques. And the symmetry supports the possibility of late insertions of rows as well as of columns.

References

- Bartels, R. H., and G.H. Golub: The Simplex Method of Linear Programming Using LU Decomposition, in: Communications ACM vol. 12 (1969), pp. 266-268, 275-278.

- Dantzig, George B.: Linear Programming and Extensions. Princepton, N.J.: Pincepton University Press 1963(a).
- Dantzig, George B.: Compact Basis Triangularisation for the Simplex Method, in: Recent Advances in Mathematical Programming. Eds. R.L. Graves and Ph. Wolfe. New York: McGraw-Hill 1963(b), pp. 125-132.
- Forrest, J.J.H., and J.A. Tomlin: Updated Triangular Factors of the Basis to Maintain Sparsity in the Product Form Simplex Method, in : Mathematical Programming vol. 2 (1972), N° 3, pp. 263-278.
- Haverly, C.A.: Additive Form of the Inverse, in: SIGMAP Bulletin Newsletter, N°12, June 1972, pp. 26-27.
- Land, A.H., and S. Powell: Fortran Codes for Mathematical Programming. London: Wiley 19073.
- Markowitz, Harry M.: The Elimination Form of the Inverse and its Application to Linear Programming, in: Management Science vol. 3 (1957), N°3, pp. 255-269.
- Muller-Merbach, Heiner: Die symmetrische revidierte Simplex-Methode der linearen Planungsrechnung, in: Elektronische Datenverarbeitung vol. 7 (1965), N° 3, pp. 105-113.
- Muller-Merbach, Heiner: Neuere Ergebnisse des Operations Research, in: Tagungsband vom III. Internationalen Kolloquium uber Anwendungen der Mathematik in den Ingenieurwissenschaften in Weimar 1965. Berlin: VEB 1966, pp. 257-266.
- Muller-Merbach, Heiner: linearen Planungsrechnung mit parametrisch veranderten Koeffizienten der Bedingungsmatrix, in: Ablauf - und Planungsforschung vol. 8 (1967), N° 3, pp. 341-354.
- Muller-Merbach, Heiner: Operations Research,3rd edition. Munchen: Vahlen 1973.
- Muller-Merbach, Heiner: Core Inverse Versus Basis Inverse: A Symmetric Concept for Revised Simplex Techniques, in: SIGMAP Bulletin N° 31, June 1982, pp. 24-26.
- Muller-Merbach, Heiner: An Obliquely Angled Branch and Bound Technique for Integer Programming, in: Angewandte Informatik vol. 25 1983(a), N° 6, pp. 252-257.
- Muller-Merbach, Heiner: Das Stufenkonzept des Algorithmenentwurfs, dargestellt am Beispiel der linearen Optimierung, in: Operations Research Proceedings 1982. Berlin, Hiedelberg, New York: Springer 1983(b), pp. 371-380.
- Orchard-Hays, William: Advanced Linear-Programming Computing Techniques. New York: McGraw-Hill 1968.
- Powell, Susan: A Development of the Product Form Algoritm for the Simplex Method Using Reduced Transformation Vectors, in: Mathematical Programming Study 4: Computational Practice in Mathematical Programming. Eds. M.L. Balinski and E. Hellerman, Amsterdam: North Holland 1075, pp 93-107.
- Stiefel, Eduard: Einfuhrung in die numerische Mathematik. Stuttgart: Teubner 1963.
- Vajda, Stephen: Mathematical Programming. Reading, Mass.: Addison Wesley 1961.
- Vajda, Stephen: Linear Programming: Algorithms and Applications. London, New York: Chapman & Hall 1981.
- Wiezorke, Bernhard: Ein wieteres Rechenschema fur die Simplexmethode, in: Unternehmensforschung vol. 5 (1961), pp. 166-171.
- Zoutendijk, G.: Mathematical Programming Methods. Amsterdam: North-Holland 1976.

Investigação Operacional, Vol.7 nº1 (Junho/1987) 27-36

LA APLICATION DE LA INVESTIGATION OPERATIVA EN SISTEMAS EDUCATIVOS

Francisco Amador Hidalgo
Manuel Cabanes Fuentes

Escuela Superior de Técnica Empresarial
Agrícola (E T E A)
Córdoba

Resumo: La experiencia de la Investigación Operativa (IO) - en sistemas sociales, aunque creciendo rápidamente - no es todavía tan rica como las aplicaciones en otros campos.

En el caso de "sistemas educativos" existe toda una gama de problemas potenciales"bien definidos", que encajan sin mayores dificultades con el enfoque metodológico de Investigación Operativa. Sin embargo, las autoridades educativas tienen que afrontar tambien otros problemas menos estructurados, "problemas difusos", en los que el número de factores relevantes es amplio y de naturaleza muy diversa (como por ejemplo, aspectos geográficos, socio-económicos financieros, administrativos, de comportamiento humano, etc.).

El artículo argumenta con un ejemplo, como en este segundo grupo de casos, los beneficios que se derivan de la aplicación de Investigación Operativa, aunque técnicamente más modestos, puedem ser en realidad muy importantes, y sobre todo, los únicos que con criterio realista son posibles.

1. El Analisis Cuantitativo en la Administración de Servicios Sociales

El enfoque que IO hace de la toma de decisiones está ampliamente establecido hoy en áreas muy diversas de la realidad, aunque en ocasiones se piensa que su potencial no está siendo totalmente desarrolado en algunas situaciones concretas (Williams 14, Simpson 11, Ackoff 1). Uno desses casos está representado por aquellas situaciones en las en que el comportamiento humano constituye un factor relevante en el análisis. La experiencia de IO en los sistemas sociales aunque creciendo rápidamente, está todavía por detras de sus aplicaciones en otros campos; y las dificultades en la planificación de los servicios sociales han sido puestas de manifesto por Chambers and Hindle (5).

Como ejemplos concretos, educación y salud son probabelemente los dos sistemas sociales más importantes que podrían (y en algunos casos en realidad ya lo hacen) beneficiarse de la aplicación de IO.

Algunos de los problemas que afectan a las aplicaciones de IO en el área de salud, presentados por Simpson (12), se pueden referir igualmente a educación. Particularmente relevante es su observación (pp. 209-10) "de que en ocasiones se pone más énfasis en modelar estructuras y manipularlas, que en determinar si se puede lograr algún avance significativo en el marco de las dificultades existentes, tanto de conceptualización como de disponibilidad de datos". La bibliografía sobre educación y planificación educativa es rica en ejemplos en ese sentido.

Se sugiere aquí, que esta observación es central a muchos de los problemas que atañen a la aplicación de IO en sistemas educativos, y especialmente relevante en aquellas situaciones, (problemas difusos) en las que el número de factores relevantes es amplio, de naturaleza muy diversa y en las quelas teorías sobre las relaciones causales en el sistema son múltiples, confusas, inconcluyentes y cuando no contradictorias.

Aplicaciones a Problemas Bien Definidos

En los sistemas educativos existe toda una gama de problemas potenciales "bien definidos", que encajen sin mayores dificultades con el enfoque metodológico de IO.

Problemas de este tipo son las estimaciones de professorado y alumnos, localización de instalaciones etc. Sobre estos problemas y su tratamiento existem diversos textos teóricos (7), revisiones bibliográficas (3) y magníficos "casos" como el realizado en Portugal en 1972, en un proyecto especial de la OCDE, con motivo del cambio del sistema Educativo. Por este motivo, no se tratan con mayor detalle en este artículo.

Problemas Difusos

Una de las razones que justifican la construcción de modelos y los convierten en un instrumento útil para la toma de diciones, es evitar el tener que experimentan con la realidad. Para ello se requiere que los modelos tengan carácter predictivo. En este artículo se entiende por problemas difusos, aquellos que debido a la complejidad de los factores que intervienen, y a la falta de conocimiento actual sobre sus interrelaciones y efectos, desafían y resisten la construcción de modelos predictivos del sistema total.

En el contexto de sistemas educativos, la assignación de recursos para la igualdad de oportunidades educativas representa uno de esos casos. Hasta qué punto es útil en estas situaciones el enfoque metodológico de IO?.

Se sugiere aquí, que la metodología de OR puede ser de utilidad incluso en situaciones como estas, en que el carácter predictivo no parece ser posible. Esto es debido, a que la racionalidad del enfoque de OR está sólo en su capacidad de síntesis (construcción del modelo), sino también en su poder de análisis (identificación de factores relevantes, classification de la realidad).

2. La Asignación de Recursos para la Igualdad de Oportunidades Educativas

El Contexto Histórico

En las últimas décadas hemos asistido a cambios dramáticos en la escala del fenómeno educativo en todos los países. Pero un análisis cuidadoso releva que, en muchos casos, esta tremenda expansión ha tenido un efecto pequeño en las disparidades que social y regionalmente existen en la participación educativa. Aunque todo el mundo está obteniendo más en términos absolutos, en términos relativos la situación es muy parecida a la anterior. Las diferencias pueden que sean mayores en unos países que en otros, pero en general, las desigualdades en participación educativa parecen ser una característica poco cambiante de las estructuras educativas.

> Qué es lo que hace de la igualdad de oportunidades en educación ser un objetivo tan difícil de alcanzar?

Existe una discusión amplia sobre este punto en la bibliografía especializada. En términos teóricos existen dos enfoques particulares del problema que merece la pena destacar debido a sus implicaciones desde la perspectiva de planificación y administración del sistema educativo.

El primero de ellos se conoce como el argumento de "la cultura de la pobreza". Como Halsey (1972) ℓ_0 resume:

> "La cultura de la pobreza insiste, en su versión más simple, en que los pobres son diferentes no fundamentalmente porque sus ingressos sean bajos sino porque han sido habituados a la pobreza y han desarrollado una sub-cultura de valores adaptados a esas condiciones que, a su vez, transmiten a sus hijos".

Las implicaciones de esta visión son claras. La capacidad de los niños de ciertos grupos socio-económicos y - àreas geográficas para beneficiarse de la educación - y en consequencia incrementar sus oportunidades en la vida - es limitada, incluso mínima. La escalera educativa que les saque de su supuesta situación de "desventaja" está, en consecuencia, psicológica e intelectualmente bloqueada. Es en este contexto en el desarrolló la estrategia de "educación compensatoria".

Un enfoque totalmente diferente ha sido propuesto recientemente por varios sociólogos e investigadores (Valentine, 1979; Pahl, 1970; Byrne and Williamson, 1973, etc.) - quienes sugieren que los patrones culturales, en la medida en que son relevantes para el éxito educativo, son, como argumenta Harris (1971), "el resultado de la respuesta de generaciones sucesivas a condiciones similares de vida social". De esta forma, una interpretación adecuada del aprovechamiento educativo no puede basarse sólamente en los "valores" de los que están en desvantaja (que son centrales en la tesis de la "cultura de la pobreza", sino que ha de tener en cuenta las condiciones sociales y económicas que son los determinantes de fondo de la situación.

La idea de un "sistema socio-espacial" es vital para la argumentación presentada por este segundo grupo de autores. Este concepto fue desarrollado por el sociólogo urbano R.E. Pahl y se puede resumir diciendo que existe un elemento espacial en la distribuición de recursos que juenga un papel relevante en la determinación de las oportunidades en la vida de los individuos y grupos. Dónde vivem los individuos o grupos puede ser un determinante fundamental de sus posibilidades de acceso a recursos deseables en la sociedad. Para estos autores, la provisión de recursos educativos realizada por las autoridades educativas, en la medida en que tal provisión varía a lo largo de las diferentes áreas geográficas y grupos socio-económicos, no solamente es un área de estudio abandonada sino, aún más, puede que tenga mayor importancia a la hora de explicar las variaciones en los patrones de desarrollo educativo, que los factores mucho más documentados asociados a la classe social, estructura social de la familia, etc.

Planificación y Control en Educación

Conceptualmente la "Planificación Educativa" no es diferente de otros tipos de planificación, aunque el entorno en el que tiene que operar presenta algunos aspectos de especial importancia.

En primer lugar, la multiplicidad de objetivos que la educación tiene que servir presenta, en sí, dificultades profundas a un análisis "racional" de políticas.

En segundo término, la educación tiene lugar en un contexto de valores. Como ha sido puesto de manifesto por Eggleston (1977):

> "El proceso de distribución de recursos en educación, está determinado por una mezcla de políticas deseadas prácticas tradicionales, supuestos, explicativas e incluso valerosas esperanzas de respuesta, en vez de por un proceso general de evaluación".

Finalmente, hay que tener en cuenta la experiencia pasada de que los planificadores no siempre han sido lo suficientemente sensibles a las distintas formas en que el proceso de planificación puede ayudar a la formulación de políticas (Eide, 1971) y con demasiada frecuencia el modelo del mundo de los planificadores tiene poco parecido con el mundo en el que los que toman las decisiones se debaten.

Este punto es fundamental a la hora de evaluar las desalentadoras conclusiones a las que están llegando organizaciones internacionales tales como la O. E. C. D. y el Banco Mundial, sobre si la expansión de acceso a la educación, en sí misma, tiende o no hacia la nivelación de las tasas de participación de los diferentes estratos sociales.

Está claro que en tipo de contexto tan amplio como los sistemas educativos, con tantos factores y facetas afectando a un mismo problema, y con opiniones contradictorias sobre relaciones funcionales básicas, es difícel la construcción de modelos predictivos.

Además el hecho de que la distribución de recursos representa, en el contexto de la igualdad de oportunidades, una condición necesaria, que según los casos puede o no ser tambien suficiente, hace que el problema de asignación equitativa de recursos sea muy difícil. Esto implica que para la asignación de recursos no será suficiente con definir objetivos de igualdad, medirlos e identificar las diferencias entre los objetivos y la realidad. La razón es que situaciones similares pueden presentar deficiencias con origen distintos; y en consecuencia distintas necesidades de recurso. Pues bien, en estas circunstancias tan poco favorables el enfoque metodológico IO puede ser de mucha utilidad. Primero, clasificando los factores fundamentales de los que no ℓ_0 son. Y segundo, diseñando sistemas de diagnosis y control que sustituyan a los modelos predictivos, hasta que el conocimiento de la función de educación permita e uso de los mismos.

El Contexto Español

La Ley General de Educación de 1970 asignaba al Ministerio de Educación la responsabilidad amplia de promocionar la educación de los españoles y requería del Gobierno un esfuerzo creciente para acabar con las diferencias regionales y sociales en materia educativa.

El Sistema Educativo Español, en estructura educativa y administrativa, y muchos de los datos publicados por el Ministerio de Educación han servindo de base real en este estudio.

Los Objetivos de la Investigación y la Metodologia

Los objetivos principales de este trabajo se pueden resumir como:

a) Es posible planificar y controlar formalmente la obtención de la igualdad de oportunidades en educación?

Se ha analizado la naturaleza de este objetivo y so pesado las posibilidades de alcanzarlo. El enfoque está basado casi enteramente en aspectos de "justicia social" en oposición a razonamientos de "utilidad".

b) Tiene la provisión de recursos educativos realmente algún efecto sobre las tasas de participación educativa de los diferentes estratos sociales y áreas geográficas, con independencia de otros factores del entorno?.

c) Sería posible integrar los resultados de a) y b) en un enfoque más sistemático y consciente de assignación de recursos basado en necesidades, y que fuese de utilidad a las autoridades educativas para la elaboración de sus políticas?.

Una característica del punto de vista adoptado a lo largo del trabajo es que no se intenta explicar formalmente cómo la planificación educativa puede interrelacionarse con otros tipos de planificación (por ejemplo, económica, de la problación, salud, etc.), ni cómo la igualdad de oportunidades se relaciona con otros objetivos posibles dentro de la planificación educativa. Esto no significa que estas tareas no se pueden o no se deban de hacer, o que sean menos importantes; pero apartarían a esta investigación de su principal área de preocupación, es decir, la assignación de recursos basada en necesidades educativas.

El Método de Análisis:

La figura 1 muestra en forma diagramática los pasos implicados en el enfoque metodológico propuesto, así como sus conexiones lógicas.

Idealmente, el objetivo sería tener tanta información y tan fidedigna como fuese necessario en cada caso. De esta forma se podría desarrollar todo tipo de análisis. Sin embargo, en la realidad es necesario trabajar dentro de las limitaciones impuestas por los datos existentes.

En el caso de planificación para igualdad de oportunidades sería conveniente disponer de datos anuales referentes a:

- Población según edad, región, zonas rurales y urbanas, y clase social.
- Número de estudiantes según edad, nivel educativo, región, zonas rurales y urbanas, y clase social.
- Porcentajes de estudiantes que se muevan a través del sistema según nivel educativo, región, zonas rurales y urbanas, y clase social.
- Recursos asignados según nivel educativo, región, zonas rurales y urbanas, y clase social.
- Distancia recorrida para ir a la escuela según nivel educativo, región, zonas rurales y urbanas, y clase social.
- Exito educativo obtenido según nivel educativo, región, zonas rurales y urbanas.

Este resume de información ideal es solamente un pequeño ejemplo del tipo de datos que se necesitarían para poder llevar a cabo un estudio completo del problema de igualdad de oportunidades.

En este trabajo la estrategia ha sido siempre la de intentar obtener la información ideal. Cuando esto no ha sido posible directamente se ha intentado obtenerla mediante la elaboración de datos más básicos. En todo momento la preocupación principal ha sido: Qué es posible hacer en términos prácticos con los datos existentes?.

La información recogida para el Sistema Educativo Español y el Andaluz se ha clasificado en cuatro apartados:

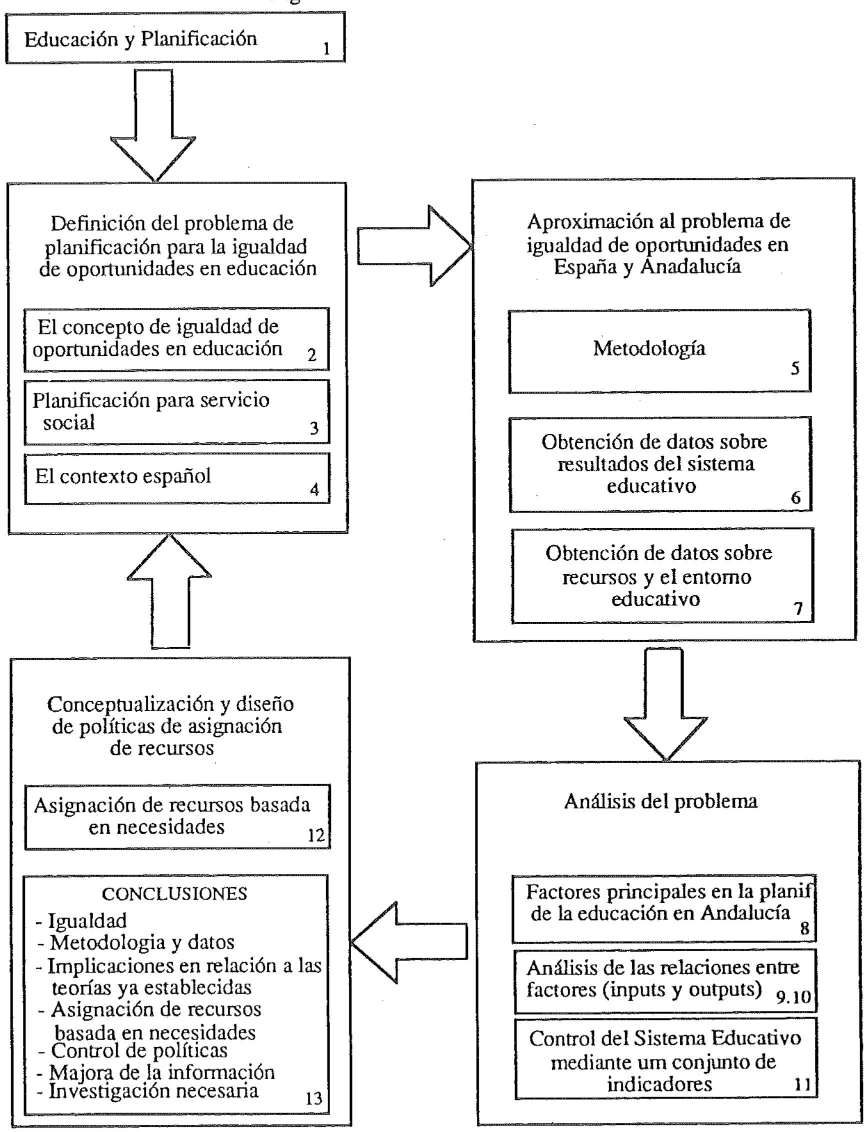
1. Problación

2. Estudiantes
3. Uso de recursos
4. Factores Socio-económicos

Los Factores Pricipales en la Planificación para la Igualdad de Oportunidades:

El conjunto de datos mencionados en el apartado anterior supone más de doscientas cincuenta variables. Este volumen de información necesita ser reducido mediante la identificación de información redundante y, cuando sea posible, la consolidación de grupos de variables semejantes en índices.

Figura 1. - Contenido del Estudio



Relaciones entre Factores Principales:

Con los resultados obtenidos del análisis de componentes principales sería posible especificar un conjunto de indicadores principales. Sin embargo, esos indicadores serían puramente descriptivos y no serían útiles para indicar los efectos de cambios, especialmente de cambios en la política de asignación de recursos. Tales efectos solamente se pueden identificar mediante un análisis causal.

El análisis de las interacciones entre factores en educación presenta algunos problemas metodológicos. Esto se debe al hecho de que las variables en la realidad escolar estan interrelacionadas y son tan interdependientes, que es muy difícil determinar sus efectos separadamente.

La situación se hace incluso más difícil porque la contribución de cualquier variable individual a la varianza total puede ser, en algunos casos, muy pequeña.

Las técnicas usadas en este trabajo para analizar la relación existente entre factores han sido la correlación parcial y la regresión múltiple.

La correlación parcial facilita al investigador una medida única de asociación describiendo la relación entre dos variables al mismo tiempo que se controlan los efectos de una más variables adicionales.

De forma adicional, los modelos de regresión desarrollados usando tasas de escolaridad como variables dependientes indican la contribución relativa de un factor específico o de un conjunto de factores.

Mediante el uso de componentes principales junto al de correlación parcial, regresión múltiole y de la revisión bibliográfica del principio de igualdad de oportunidades, se ha definido un conjunto de indicadores destinados a identificar necesidades, definir objetivos y controçar el progresso hacia esos objetivos.

El Control de un Sistema Mediante un Conjunto de Indicadores Principales:

Una vez se ha definido un conjunto de indicadores referentes a los aspectos relevantes de un sistema educativo (especialmente en relación con la igualdad de oprotunidades), hay dos líneas de acción que es necesario seguir.

1) Integrar el conjunto de indicadores en un marco de planificación capaz de detectar rápidamente las desviaciones entre sus "valores objetivos" y sus "valores reales".
2) Explorar las implicaciones de 1), conjuntamente con las posibles relaciones causales existentes entre factores que miden outputs educativos y aquellos otros relacionados a los inputs y el entorno, con respecto a la política de asignación de recursos.

El conjunto de indicadores está diseñado siguiendo el concepto de "conos de resolución" de BEER (1967), permitiendo al planificador observar el sistema educativo a distintos niveles de detalle y rastrear (incrementando el nivel de detalle a medida que va siendo necesario) el origen de problemas detectados a altos niveles de agregación.

El conjunto de indicadores se puede aplicar igualmente a nivel nacional, provincial o municipal, y contituye la base para la definición de necesidades educativas, objetivos y políticas.

Políticas de Asignación de Recursos Basadas en Necesidades:

Existen dos aspectos que ninguna política de asignación de recursos en educación puede ignorar si es que ha de tener algún efecto sobre los objetivos educativos, especialmente sobre el objetivo de igualdad de oportunidades. Esos aspectos son, la identificación de necesidades y la traducción de esas necesidades en términos del tipo de recursos que las autoridades educativas pueden utilizar en su labor de administradores del sistema.

Algunos Aspectos Básicos para Análisis y Resultados

Desigualdades en la participación educativa existen en muchos países. En el caso del sistema Educativo Español, se observa que existen diferencias en la participación educativa a varios niveles - entre provincias, áreas rurales y urbanas y grupos sociales -.

Se sugiere que en el caso del Sistema Educativo Español, y probablemente en muchos otros casos, un enfoque global y genérico del problema de la igualdad de oportunidades no es posible que tenga éxito. Esto es porque, a ese nivel de agregación, hay demasiados efectos que tienen lugar al mismo tiempo para que sea posible la identificación de soluciones relevantes. En su lugar, se sugiere que el enfoque del problema de la igualdad de oportunidades se debería de basar en la identificación de las necesidades y condiciones locales.

Esta sugerencia implica algunos puntos básicos:

a) Una definición operacional de la igualdad de oportunidades educativas y cómo medirla.

b) La definición de un conjunto de indicadores capaz de identificar los aspectos relevantes de un sistema educativo - especialmente aquellos aspectos relativos a la necesidad.

c) Un sistema de seguimiento y control capaz de identificar áreas problemáticas y las condiciones locales.

La tasa de escolarización por grupos de edad y nivel educativo se ha mostrado que son una forma eficiente y controlar la participación educativa en un sistema educativo. Las tasas de escolarización conjuntamente con algunas medidas del nivel de provisión educativa y de los factores ambientales constituyen el conjunto de indicadores principales definido después de realizar varios análisis de componentes principales, correlación parcial y regresión múltiple sobre un gran número de variables básicas.

El uso del análisis de componentes principales en la definición de un conjunto de indicadores principales, como complemento a otros enfoques posibles, es recomendable por varias razones. En primer lugar, y dado que es un análisis enteramente formal basada en medidas cuantitativas de los inputs y salidas del sistema, sus resultados están libres de juicios subjetivos a obtener conclusiones contrarias a partir de esencialmente la misma información.

Los resultados del análisis de componentes principales facilitarían por sí solos la definición de un conjunto de indicadores principales, pero estos serían principalmente descriptivos y no indicarían los efectos de cambios - especialmente cambios en la provisión de educación -. Tales efectos se pueden identificar sólamente mediante análisis causal.

Los análisis de correlación parcial y regresión múltiple sugieren que la provisión de instalaciones, los ingresos y el nivel educativo familiar son, todos, factores con efectos causales potenciales importantes sobre la participación educativa.

Con respecto a la polémica bibliográfica sobre la importancia de la provisión de instalaciones, el análisis del caso español (probablemente también aplicable a otros casos) muestra que tres efectos principales pueden tener lugar al mismo tiempo:

a) La provisión de instalaciones tiene una correlación independiente significativa con algunas de las tasas de escolarización. En estes casos, la participación educativa no puede explicarse solamente en términos de los ingresos y un entorno favorable en general. En consecuencia, se sugiere aquí que la explicación más probable es que la provisión tiene, en sí misma, un efecto positivo sobre la participación educativa.

b) Los efectos de la provisión de instalaciones, el nivel educativo familiar y los ingresos son interdependientes.

c) La provisión es baja cuando el nivel educativo familiar y los ingresos son bajos - esto es, por supuesto, inconsciente con la prioridad de objetivos que las autoridades educativas proclaman normalmente.

Las regresiones que se han usado para explicar la participación educativa representan una visión a medio y largo plazo de los factores principales relacionados a la participación educativa y de su importancia relativa. Pero estas ninguna ligazón.

La Asignación de Recursos Basada en Necesidades:

Se sugiere como una de las conclusiones de esta investigación que el enfoque del problema de asignación de recurso debería de basarse en un proceso de tres etapas:

a) El diagnóstico de áreas problemáticas y la identificación de las condiciones locales.

b) La agregación de las necesidades clasificadas en el nivel más bajo del análisis (nivel municipal) para dar lugar a una medida de necesidad en los niveles más altos del sistema (provincias y regiones).

c) La asignación de recursos basada en necesidades

- El controle del sistema. La fase de diagnóstico

Toda la discusión presentada hasta ahora sería de poca utilidad a menos que exista una forma de aislar áreas problemáticas y de identificar las condiciones locales, de forma que el proceso de asignación de recursos pueda ser dirigido hacia necesidades específicas.

Esto se consigue mediante la aplicación metódica de un sistema de seguimiento y control con dos ejes de desagregación. El primer eje es geográfico - se pueden identificar de esta áreas de problemas a nivel nacional, provincial y de distrito (o municipal). En cada nivel geográfico un segundo eje de desagregación guía el análisis desde los resultados educativos (es decir, tasas de escolarización por grupos de edad y niveles educativos) hasta los inputs educativos y factores ambientales.

Este enfoque tiene dos ventajas principales:

a) Reduce el número de dimensiones del problema que se han de tener en cuenta a la vez en cualquier momento.
b) Favorece los aspectos de implementación ya que algunas de las variables se reducen a variables locales.

- La medida de las necesidades

Dado que las provincias son la suma de sus distritos, y puede que haya grandes variaciones entre distritos dentro de una misma provincia, la forma lógica de medir las necesidades y de formular los objetivos para una provincia es mediante la agregación de las necesidades de los distritos. Las provincias estimarían entonces sus propias necesidades mediante el análisis y agregación de los distritos individuales.

La medida de necesidad que se obtiene al aplicar la tasa de escolarización 'objetivo' a poblaciones en el grupo de edad teórica para cada nivel educativo, tiene que clasificarse por fuentes específicas de necesidad, tales como escuelas hogar, puestos de comedor, transporte, etc.

- El proceso de asignación

El proceso de asignación de recursos opera en dirección opuesta a la medida de necesidades. Los recursos se asignan primero a las provincias en proporción a su medida de necesidad, y una vez que las disponibilidades provinciales de recursos quedan estabelecidas, entoces se distribuyen a los distritos siguiendo el mismo principio.

- Las ventajas del enfoque

La descomposición del problema de la igualdad de oportunidades educativas en áreas más pequeñas, mediante el eje geográfico de desagregación, es un aspecto importante del enfoque. Tiene no sólo la ventaja de tener que enfrentar problemas más pequeños en cualquier fase del análisis, sino también la de tener en cuenta la estructura administrativa del sistema.

La ventaja de enfrentar problemas más pequeños se ve claramente en el caso de las áreas rurales. Una vez que se toman medidas correctivas a nivel de distrito (o municipal), el problema rural queda resuelto casi por definición.

Se sugiere que este enfoque representa una visión realista del tipo de progreso que se puede conseguir en la solución del problema de la igualdad de oportunidades en educación. Planes globales, técnicamente impecables pero que no tienen en cuenta las características de los sistemas humanos, y que no reflejan el mundo en el que las autoridades educativas se mueven, no son muy útiles.

Finalmente, cualesquiera que sean los detalles del sistema de asignación específico adoptado en los diferentes niveles de la estructura administrativa, hay dos puntos principales que deberían de tenerse en cuenta:

a) Se debería de utilizar un enfoque consistente siempre que sea posible - es decir,en cada situación, identificar las necesidades relativas, comparar las necesidades con la provisión existente y sugerir un plan práctico para pasar del patrón existente a nuevos patrones de asignación.
b) El proceso debería de ser lo suficientemente flexible como para dar a las autoridades educativas la oportunidad de modificar un plan para tener en cuenta otros criterios que el de necesidad.

Un Ejemplo para la Asignación de Capital :

Las figuras 2 y 3 muestran cómo un procedimiento de asignación basado en necesidades da, en el caso de España, resultados muy diferentes de los de la política actual tanto a nivel provincial como de distrito.

Todos estos datos se refieren a la provisión de puestos de educación secundaria (BUP+COU) en el período 1975-81.

La figura 2 refleja la relación al objetivo de asignación para cada provincia. Los objetivos de cada provincia para 1981 se han establecido suponiendo una tasa de escolarización objetivo 'general' para el grupo de edad teórica en 1981 (14-17 años) igual al valor provincial más alto en 1975.

De esta forma, la tasa de escolarización provincial más alta se usa como el objetivo teórico para alcanzar igualdad. La diferencia entre el número existente de puestos en cada provincia en 1975 y su objetivo para 1981, sería su medida de necesidad.

Los objetivos se comparan entonces con el número real de puestos existentes en cada provincia en 1981. Esto es, por supuestos, una comparición simplificada. No se han tenido en cuenta, por ejemplo, factores de crecimiento potencial. Sin embargo, incluso a este nivel de aproximación se identifican diferencias entre objetivos y asignaciones efectivas que son de difícil justificación.

La figura 3 compara la distribución porcentual 'objetivo' de los recursos de capital disponibles, con los valores efectivos. En este caso las provincias se han representado en orden ascendente de su participación porcentual en recursos disponibles.

A nivel de distrito, las figuras 4 y 5 muestran la misma comparición. Los objetivos a nivel de distrito se han calculado de la misma forma que para las provincias. Sin embargo, en este caso, las tasas de escolarización se han medido sobre la población total del distrito en 1981 en vez de los grupos teóricos de edad.

Figura 2. - Ilustración de la relación al objetivo de asignación de capital para cada provincia (1975/81)

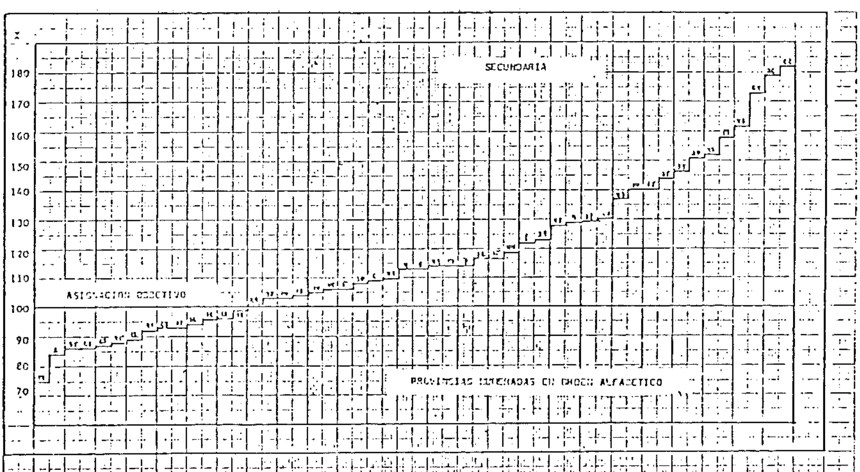


Figura 3. - Asignación porcentual de recursos de capital a las provincias: Asignación Objetivo y Real (1975/81)

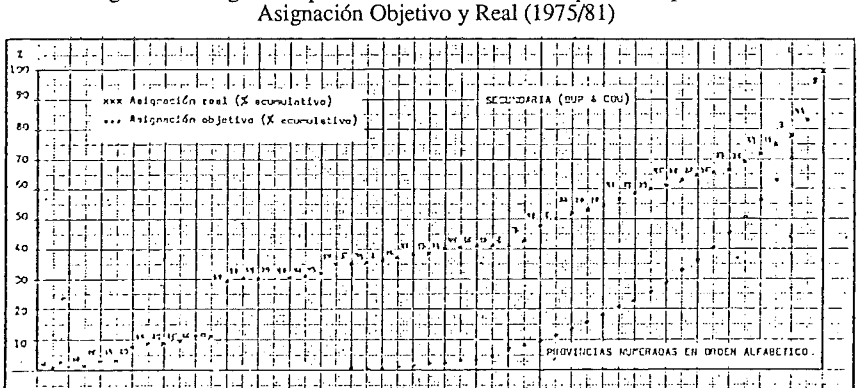


Figura 4. - Asignación porcentual de recursos de capital a los distritos: Asignación Objetivo y Real. Córdoba (1975/81)

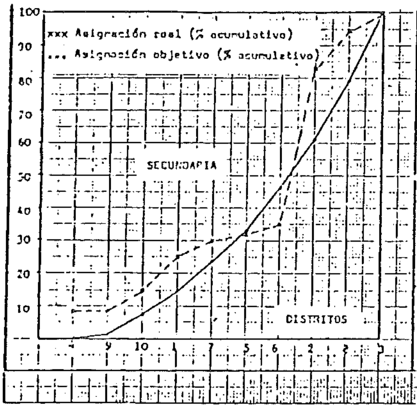
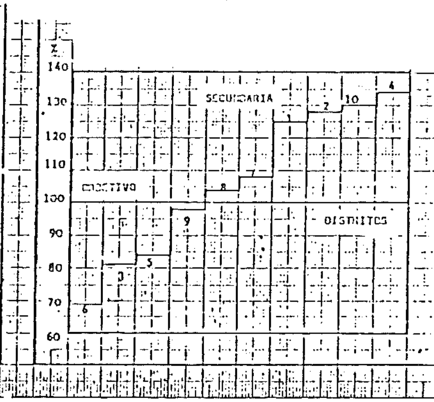


Figura 5. - Ilustración de la relación al objetivo de asignación de capital para cada distrito. Córdoba 1975/81.



Referências

(1) ACKOFF, R. L. - *"The future of Operacional Research is Past"*, Journal of the Operational Research Society, Vol. 30, nº 2, pp. 93-104, (1979).
(2) BEER, S. - *"Management Science : The Business Use of Operation Research"*, Aldus, London, (1976).
(3) BLAUG, M. - *"An Introduction to Economics of Education"*, the Penguin Press, Harmondsworth, (1970).
(4) BYRNE, D., FLETCHER, B. and WILLIAMSON, B. - *"Resource Allocation : A Theoretical Viewpoint"*, in The Urban Context, the Open University, E351, Block 2, the Open University Press, Porstmouth, (1974).
(5) CHAMBERS, M. L. and HINDLE, A. - *"Problems in Planning for Service"*, in Reader in Operations Research for Libraries, (Brophy, P. et al, Ed.), Information Services Englewood, Colorado, (1976).
(6) EGGLESTON, J. - *"The Ecology of the School"*, Methuen, London, (1977).
(7) FOWLER, G. - *"Decision Making in British Education"*, The Open University, London,(1973).
(8) HALSEY, A. H. ed. - *" Educational Priority"*, HMSO, London (1972).
(9) HARRIS, M. - *"Culture, Man and Nature"*, Crowell, New York, (1971).
(10) PAHL, R. E. - *"Whose City ?"*. Longman, Londen, (1970).
(11) SIMPSON, M. G. - *"Production"*, Paper presented at a Discusion Conference sponsored by the OR department of the University of Lancaster, Chester, (1976).
(12) SIMPSON, M. G. - *"Health"*, Opl. Res. Q., Vol. 27, nº 1, pp. 209-219, (1976).
(13) VALENTINE, C. A. - *"Culture and Poverty: Critique and Counter - Proposals"*, Chicago University Press, London, (1968).
(14) WILLIAMS, G. L. - *"Introduction and Summary of the Discussions"*, in Efficiency in Resource Utilization in Education, pp. 11-20, OCDE, París, (1969).

Investigação Operacional, Vol.7 nº1 (Junho/1987) 37 - 45

MODELO DE PROGRAMAÇÃO DA MANUTENÇÃO PREVENTIVA DE GRUPOS TÉRMICOS INTEGRADOS NUM SISTEMA ELECTROPRODUTOR HÍDRICO-TÉRMICO

Abílio Seca Teixeira

Direcção Central de Planeamento (EDP)

Resumo: Esta comunicação tem por objectivo a apresentação de uma possível metodologia de programação das acções de manutenção preventiva de grupos térmicos produtores de energia eléctrica, elemento importante para a simulação da exploração optimizada de um sistema electroprodutor, num determinado estádio de evolução.

O modelo desenvolvido procura determinar o esquema geral de manutenção programada dos diversos grupos térmicos, visando minimizar os encargos de exploração, resultantes da operação do sistema electroprodutor, tendo em conta restrições de ordem vária, associadas ao funcionamento do parque térmico.

O objectivo a atingir, escolha do melhor conjunto de intervalos de tempo na perspectiva apontada, para a realização da manutenção dos grupos térmicos, traduz-se num problema de tipo combinatório, de dimensões muito elevadas, que se procura resolver formulando-o como um problema de programação linear em variáveis mistas.

Abstract: The main goal of this paper is the presentation of one model developed to define an optimal maintenance scheduling policy for the thermal units in a hydro-thermal generating system.

The problem to solve deals with the definition of the starting dates for the maintenance actions for every thermal power unit, aiming at the minimization of the expected value of the variable operating costs, taking into account all the relevant constraint and random conditions that affect the operation of the power generation system.

This large combinatorial problem is solved by utilization of linear programming with mixed variables.

Keywords: Maintenance, Crews, Heuristic, Mixed Variables, Sequential Minimization.

1. Introdução

Todo o sistema produtor está destinado a exercer uma função determinada de satisfação duma procura, à qual, em geral, se associa um custo unitário de produção, que se procura minimizar.

Por outro lado, qualquer sistema produtor está sujeito a avarias de diversos tipos, com carácter aleatório no tempo, que influenciam desfavoravelmente a qualidade e o custo dos serviços por ele prestados. Para que com uma dada probabilidade de sucesso, ele possa corresponder às solicitações exteriores, é em geral necessário proceder a acções de manutenção preventiva das suas unidades componentes, tendentes a reduzir a frequência das avarias fortuitas. A estas acções de manutenção associam-se, porém, certos cargos que importa determinar, de forma que o custo global de produção resulte mínimo.

Haverá portanto interesses em determinar, dentro de condições técnico-económicas de garantia e segurança, uma política de manutenção que conduza a esse custo mínimo. Torna-se portanto necessário definir os critérios a minimizar e determinar as condições consideradas impressionáveis para a garantia de uma manutenção eficaz.

A fase de planeamento de unidade de produção tem em conta variados indicadores, fornecidos por modelos matemáticos mais ou menos elaborados. Esses modelos tentam simular a exploração optimizada de um sistema produtor-consumidor, suposto conhecido nos modelos deterministicos, ou provável nos estocásticos.

A noção de custo mínimo é relativa num modelo determinista, uma vez que depende das condições em que as unidades de produção foram postas à disposição do modelo, para simulação. Por outras palavras, o custo mínimo dependerá necessariamente da política de manutenção adoptada para o sistema produtor. Surge assim a necessidade de introduzir na fase de planeamento, e a par de outros, um modelo que estabeleça uma política "óptima" de manutenção para o sistema.

Poder-se-á pensar que, no planeamento a médio e longo prazo, pouco importa a distribuição das acções de manutenção no tempo, bastando apenas considerar, em face de estatísticas e características de fabrico, uma distribuição uniforme daquelas acções.

Contudo, no caso de um sistema produtor de energia eléctrica hidrotérmico, aquele que é analisado nesta comunicação, a experiência tem mostrado que os resultados fornecidos pelo modelo de exploração vinham significativamente diferentes para distribuições igualmente diferentes das acções de manutenção, num mesmo intervalo de tempo.

A partir desta constatação fomos levados à formulação do modelo a seguir apresentado, de modo que os indicadores e resultados obtidos a partir da exploração optimizada do sistema não viessem distorcidos por uma política de manutenção menos adequada.

O modelo foi concebido apenas para o subsistema térmico, porque, relativamente aos aproveitamentos hidroeléctricos,e entre outras razões:

- as ligações hidráulicas entre aproveitamentos duma cascata e o partido que daí se pode tirar quando se procede a acções de manutenção sobre os seus aproveitamentos;
- o carácter não periódico de algumas acções de manutenção de longa duração sobre alguns dos orgãos que os compõem;
- a constatação de uma maior fiabilidade dos componentes das centrais hidroeléctricas, relativamente aos equivalentes térmicos;
- a capacidade de adaptação dos fios-de-água aos caudais afluentes, em geral devido a um sobreequipamento, que permite a execução de acções de manutenção preventiva em grupos cuja utilização não é justificável para aqueles caudais;
- a possibilidade de se proceder a acções de manutenção nos aproveitamentos de albufeira quando a gestão óptima das suas reservas impõe o armazenamento.

levaram a concluir que, numa primeira fase, as acções de manutenção preventiva referentes aos componentes do subsistema hidroeléctrico poderiam admitir-se uniformemente distribuidas no tempo.

2. O Modelo "Manutenção"

2.1 Condições mínimas a que deverá obedecer o modelo

Já se referiu a necessidade de fornecer ao modelo de exploração os dados correspondentes a uma política de manutenção adequada, ou, por outras palavras, as potencias térmicas disponíveis em cada unidade de tempo. Numa breve referência ao modelo de exploração utilizado, suponhamos conhecidos para um dado período de tempo $[0, T]$:

- a função $D : [0, T] \rightarrow \mathbb{R}^+$, definidora da procura de potencia;
- a composição do sistema electroprodutor

$$S : [0, T] \rightarrow \mathbb{R}^{np} \times \mathbb{R}^{nH}$$
$$t \rightarrow (P(t), H(t))$$

  que a cada instante t associa as componentes térmica P(t) e hidroeléctrica H(t);
- os custos variáveis de produção das unidades térmicas

$$C_k : [0, T] \times [0, \bar{P}_k] \rightarrow \mathbb{R}$$
$$(t, p_k) \rightarrow C_k(t, p_k)$$

  onde $\bar{P}_k$ representa a potência máxima disponível da unidade k.
- as afluências aos aproveitamentos hidroelétricos

$$w_{n,r} : [0, T] \rightarrow \mathbb{R}$$
$$t \rightarrow w_{n,r}(t)$$

- a rede de transporte
- a função M que define a política de manutenção das unidades produtoras.

O objectivo consiste em minimizar os custos totais de produção

$$\int_0^T \sum_k C_k(t, p_k(t))\, dt \tag{2.1}$$

de modo a serem satisfeitas as restrições,

$$\begin{aligned} P(t) + H(t) &\geq D(t) \\ H(t) &\in \mathcal{H}(t) \\ P(t) &\in \mathcal{P}(t) \end{aligned} \tag{2.2}$$

onde $\mathcal{H}(t)$ e $\mathcal{P}(t)$ representam os domínios das variáveis H(t) e P(t), respectivamente.

Sendo o modelo de exploração determinista no que respeita às funções **S**, **D** e **M**, é evidente que os resultados virão influenciados pelos seus valores. O carácter não determinista da função **w** é eliminado simulando com uma série hidrológica do passado e tomando valores médios.

A partir do momento em que se verificou a importância do papel da função **M** do modelo de exploração concluiu-se que o modelo de manutenção deveria obedecer no mínimo às seguintes características:

- aos grupos térmicos são atribuidos dois modos de funcionamento:
  1 - totalmente disponíveis (modo B)
  2 - totalmente indisponíveis (modo H)
- a probabilidade de uma avaria fortuita é constante ao longo do período de funcionamento.
- as acções de manutenção possíveis são:
  1 - ausência de acção ou espera
  2 - substituição do grupo térmico
- as restrições impostas são:
  1 - continuidade das acções de manutenção
  2 - satisfação de condições técnicas para a realização da tarefa,
  3 - satisfação da procura de potência.
- o custo a minimizar será dado pelos encargos em combustível associados à exploração dos grupos térmicos, o que para um estádio definido, significa minimizar os custos unitários de produção.

2.2 Descrição do modelo

2.2.1 Noções Gerais

Antes de iniciarmos a descrição pormenorizada da formulação matemática do modelo convem explicitar algumas das noções utilizadas, tais como: *custo de substituição, continuidade da acção de manutenção e restrições técnicas à acção.*

Admitamos a existência de um sistema com K+1 grupos térmicos $\{G_1, G_2, \ldots, G_{K+1}\}$ ordenados por ordem crescente dos seus custos variáveis. G_{K+1} representará o grupo de custos mais elevados e 100% fiável, não estando portanto sujeito a qualquer tipo de indisponibilidade (é equivalente a um centro penalizador de energia não fornecida pelo sistema).

Se o grupo G_k fizer manutenção preventiva no intervalo $[\, t_1, t_2\,] \subset [\, 0, T\,]$, e considerando:

- $C^0(k)$ o mínimo custo de produção obtido em $[\, 0, T\,]$, admitindo que G_k não entra em manutenção, e que o seu modo de funcionamento é $\underline{B}$, $\forall t \ \varepsilon\ [\, 0, T\,]$,
- $C^1(k, t_1, t_2)$ o mínimo custo de produção obtido em $[\, 0, T\,]$ quando G_k faz manutenção em $[\, t_1, t_2\,]$ e que $\forall t\ \varepsilon\ [\, t_1, t_2\,]$ o seu modo de funcionamento é $\underline{B}$.

o *custo de substituição* do grupo G_k no intervalo de tempo $[\, t_1, t_2\,]$ para uma determinada composição do sistema será:

$$C(k, t_1, t_2) = C^1(k, t_1, t_2) - C^0(k) \tag{2.3}$$

Os dados actualmente disponíveis dão-nos ideia da diversidade do tempo necessário à realização das acções de manutenção preventiva, de 3 a 12 semanas conforme o tipo de grupo térmico.

Para garantir que, uma vez iniciada uma acção de manutenção, ela se realize continuamente até ao seu final, o modelo considera o que se designa por *continuidade da acção de manutenção.*

Finalmente refira-se que um determinado intervalo de tempo pode ser óptimo para a realização da manutenção preventiva de um grupo G_k mas, devido à insuficiência de meios técnicos ou humanos, aquela acção não possa ser realizada no intervalo escolhido.

Este é o tipo de restrições que se designa por *restrições técnicas à acção de manutenção* e que está implementado no modelo pelo conceito de "equipas de manutenção".

2.2.2 Formulação matemática

Considere-se um sistema electroprodutor com K+1 grupos térmicos, N centrais hidroeléctricas e um diagrama de cargas a satisfazer D(t), classificado em J postos horários.

Admitamos as seguintes hipóteses

H1 - O intervalo de estudo das acções de manutenção é limitado $[\, 0, T\,]$, e está discretizado em T pontos $\{1, 2, \ldots, T-1, T\}$;

H2 - Conhece-se o valor do diagrama de cargas $D_{j,t}$, para o posto horário j e intervalo de tempo t ;

H3 - Conhece-se a contribuição em potência $h_{r,n,j,t}$ da central hidroeléctrica n, no posto horário j e intervalo de tempo t, na ocorrência do regime hidrológico r ;

H4 - Só se admite uma acção de manutenção sobre cada grupo térmico durante o intervalo de estudo $[\, 0, T\,]$;

H5 - Para a descrição completa do estado de um grupo térmico G_k são suficientes as seguintes variáveis e funções:

H5.1 - independentes do tempo

p_k^* - potência instalada,

s_k - tempo necessário à realização da manutenção preventiva,

t_k - instante em que tem início essa manutenção.

H5.2 - dependentes do tempo

$i_{k,t}$ - indicador do estado de exploração

=1 - em serviço

=0 - fora de serviço, por razões exteriores à manutenção.

$m_{k,t}$ - indicador da acção de manutenção

=0 - em serviço

=1 - fora de serviço, para manutenção.

$f_{k,t}$ - taxa de indisponibilidade forçada

$\alpha_{k,j}$ - coeficiente de limitação horária de potência no posto horário j $(\alpha_{k,j}\ \varepsilon\ [\,0, 1\,])$

$p^1_{k,t}$ - potência disponível

$p_{r,k,j,t}$ - potência efectivamente posta em serviço no posto horário j do intervalo de tempo t, na ocorrência do regime hidrológico r.

H6 - Considere-se uma partição dos recursos para realização de acções de manutenção, E, em L conjuntos E_ℓ, $\ell = 1, 2,\ldots, L$ que designaremos por tipo de equipas de manutenção. Cada uma delas tem uma capacidade de resposta $e_{\ell,t}$ função do tempo. No nosso caso significa o número máximo de grupos térmicos que o tipo de equipa E_ℓ poderá reparar simultaneamente no intervalo de tempo t.

O problema que se pretende resolver pode ser formulado do modo seguinte:

$$\text{minimizar} \quad \sum_r pb_r \sum_{k,j,t} \varnothing(r, k, j, t, p_{r,k,j,t}) \tag{2.4}$$

sendo pb_r a probabilidade de ocorrência associada a cada uma dos r regimes hidrológicos considerados,

sujeito a:

R1 - garantia de satisfação dos consumos

$$\sum_k p_{r,k,j,t} + \sum_n h_{r,n,j,t} - D_{j,t} = 0,\ \forall r,j,t \tag{2.5}$$

R2 - limitação da capacidade de produção térmica

$$0 \leq p_{r,k,j,t} \leq p^1_{k,t} \times \alpha_{k,j}\ , \forall r,k,j,t \tag{2.6}$$

R3 - potência máxima disponível por grupo térmico

$$p^1_{k,t} = (1 - m_{k,t}) \times i_{k,t} \times (1 - f_{k,t}) \times p^*_k,\ \forall k,t \tag{2.7}$$

R4 - uma acção de manutenção por grupo

$$\sum_{t=1}^{T} m_{k,t} = s_k\ , \forall k \tag{2.8}$$

R5 - garantia de continuidade no tempo de uma acção de manutenção

$$\sum_{t=t_k}^{t=t_k+s_k-1} m_{k,t} = s_k\ , \forall k \tag{2.9}$$

R6 - restrições à utilização dos recursos

$$\sum_{k \varepsilon E_\ell} m_{k,t} \leq e_{\ell,t}\ , \forall \ell,k \tag{2.10}$$

R7 - restrições ao período de manutenção

$$t^0_k < t_k < t^1_k\ , \forall k \tag{2.11}$$

com

$$t^1_k - t^0_k \geq s_k - 1$$

R8 - restrições às acções de manutenção

$$m_{k,t}\ \varepsilon\ \{0, 1\}\ , \forall k, t \tag{2.12}$$

2.2.3 Resolução do problema

Estamos perante um problema de programação matemática nas variáveis reais $p_{r,k,j,t}$ e nas variáveis inteiras $m_{k,t}$ e t_k. Na sua resolução as maiores dificuldades são introduzidas pela variável t_k e pela restrição 2.9.

Uma vez conhecido t_k, obedecendo à restrição 2.11, a determinação dos $m_{k,t}$ que satisfaçam 2.8, 2.9 e 2.12 é imediata. De facto,

$$\begin{aligned} m_{k,t} &= 0\ , \ \forall\ t < t_k \\ m_{k,t} &= 0\ , \ \forall\ t > t_k + s_k - 1 \\ m_{k,t} &= 1\ , \ \forall\ t_k \leq t \leq t_k + s_k - 1 \end{aligned} \tag{2.13}$$

o que conduz à resolução do problema da melhor colocação dos grupos térmicos no diagrama de cargas (restrições 2.5 e 2.6).

A determinação unívoca de $m_{k,t}$ pelo conhecimento de t_k permite que o problema seja formulado apenas nas variáveis t_k e $p_{r,k,j,t}$. Como para cada vector $(t_1,\ldots, t_k)$ existe uma solução única em termos de $p_{r,k,j,t}$ podemos afirmar que existe uma função

$$\varnothing^* : \{1,\ldots, T\}^k \rightarrow \mathbb{R}^+$$
$$(t_1,\ldots, t_k) \rightarrow \varnothing^* (t_1,\ldots, t_k)$$

cujo mínimo é a solução do problema original.

É evidente que se trata de um problema combinatório de elevadas dimensões..Para ultrapassar esta dificuldade admitiu-se numa primeira fase um *método heurístico* que, quando aplicado a este modelo, tem conduzido a resultados que nos permitem aceitá-lo. Assim, é dada prioridade de manutenção preventiva aos grupos térmicos que conduzam à custos de substituição mais elevados. Por aplicação deste princípio obtem-se um vector $(t_1, t_2,\ldots, t_k)$ de instantes de entrada em manutenção dos diferentes grupos térmicos, que não garantindo ser aquele que efectivamente minimiza $\varnothing^*$, representará, nas condições indicadas, uma boa aproximação do óptimo. Numa segunda fase, partindo do vector inicialmente obtido procura-se optimizar o encargo em combustível resultante da aplicação desse programa de manutenção preventiva, por recurso a um *processo de minimização sequencial.*

Em seguida faz-se uma breve referência ao algoritmo geral utilizado.

1. Seja Q^K o conjunto dos K grupos que devem entrar em manutenção preventiva no intervalo de tempo $[0, T]$.
   Seja $n = K$
2. Determinar o grupo que entra em manutenção (heurística)
   2.1 Para cada grupo k em Q^n calcular
   - os custos de substituição $C(k,t, t+s_k-1)$ em cada instante t, satisfazendo as restrições 2.8 a 2.12.
   - o intervalo $[t_k^0, t_k^0 + s_k - 1]$, onde o custo de substituição é mínimo para o grupo k.

   2.2 O grupo que entra em manutenção no intervalo $[t_z^0, t_z^0 + s_z - 1]$ será aquele cuja manutenção preventiva conduz a custos de substituição mais elevados
   $$C(z, t_z^0, t_z^0 + s_z - 1) = \max_{k \varepsilon Q^n} C^m(k, t_k^0, t_k^0 + s_k - 1)$$
3. Retirar o grupo z ao conjunto Q^n
   $$Q^{n-1} = Q^n - \{z\}$$
4. Fazer $n = n - 1$. Se $n > 0$ voltar a 2.
   Se $n = 0$ ficou determinado o vector $(t_1^0,\ldots, t_K^0)$ de instantes de entrada em manutenção dos K grupos térmicos, a partir do qual se obtem $\varnothing_k^0 (t_1^0,\ldots, t_K^0)$ encargo em combustível associado a esse vector.
5. Minimização sequencial
   Seja i um contador de iterações. Com $i = 1$
   5.1 Para $k = 1$, K calcular
   t_k^i tal que
   $$\min_{t_k \varepsilon \Omega_k^i} \varnothing_k^i (t_k; t_1^i,\ldots, t_{k-1}^i, t_{k+1}^{i-1},\ldots, t_K^{i-1}) =$$
   $$= \varnothing_k^i (t_k^i; t_1^i,\ldots, t_{k-1}^i, t_{k+1}^{i-1},\ldots, t_K^{i-1})$$
   onde
   a) Ω_k^i é o domínio admissível de t_k quando os grupos 1,..., k-1, k+1,..., K entraram em manutenção nos instantes de tempo $t_1^i,\ldots, t_{k-1}^i, t_{k+1}^{i-1},\ldots, t_K^{i-1}$.
   b) $\varnothing_k^i (t_k; t_1^i,\ldots, t_{k-1}^i, t_{k+1}^{i-1},\ldots, t_K^{i-1})$ é o encargo mínimo em

combustível quando os grupos 1,..., k-1, k, k+1,..., K entraram em manutenção nos instantes $t_1^i,\ldots, t_{k-1}^i, t_k, t_{k+1}^{i-1},\ldots, t_K^{i-1}$.

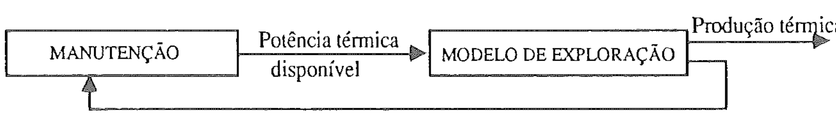
5.2 Se $| \varnothing_k^i (t_1^i,\ldots, t_K^i) - \varnothing_k^{i-1} (t_1^{i-1},\ldots, t_K^{i-1}) | < \varepsilon$ terminar.
Caso contrário fazer i = i + 1 e voltar a 5.1.

3. Interligação do Modelo Manutenção com o Modelo de Exploração

Embora o modelo MANUTENÇÃO procure minimizar os custos de exploração do parque térmico, há que ter em conta o facto de se admitir conhecida, por regime hidrológico, a contribuição hidroeléctrica, obtida por simulação com o modelo de exploração do sistema electroprodutor, para uma política de manutenção previamente fixada.

Coloca-se assim a questão de saber se as modificações na distribuição temporal das acções de manutenção dos grupos térmicos irão ou não influenciar significativamente a gestão das albufeiras do sistema. Os resultados que se têm obtido com o modelo de exploração tem mostrado que as alterações não são significativas.

Contudo, como a análise dos resultados fornecidos pelo MANUTENÇÃO não nos permite responder directamente à questão levantada torna-se necessário iterar entre os dois modelos até que os valores mais característicos se mantenham praticamente entre as duas iterações sucessivas



Assim, podemos interpretar a política de manutenção dos grupos térmicos como a procura de uma sugestão optimizada da potência térmica disponível.

4. Exemplo de Aplicação

No QUADRO I apresentam-se alguns dos resultados obtidos para a exploração simulada do estádio 1986, admitindo a acorrência de uma regime hidrológico cujas produções hidroeléctricas representam a média das produções dos regimes considerados (1964 a 1975), para três diferentes esquemas de manutenção programada do parque térmico existente.

O esquema de manutenção de referência, U, traduz-se pela distribuição uniforme, ao longo do ano, das acções de manutenção programada dos diferentes grupos térmicos. Os esquemas de manutenção designados por CR e SR (ver QUADROS II e III, respectivamente) foram obtidos por aplicação do modelo MANUTENÇÃO. No primeiro caso admitiu-se que apenas uma grupo de cada vez, por central, podis realizar manutenção preventiva, com excepção da central do Carregado em que poderiam estar simultaneamente em manutenção dois grupos durante os 1º e 4º trimestres. No caso do esquema SR considerou-se que nas centrais de Setúbal, Carregado e Tapada do Outeiro poderiam estar simultaneamente em manutenção ao longo do ano dois grupos, número esse que podia ir até três para o caso das centrais com turbinas a gasóleo.

Analisando os resultados indicados no QUADRO I é de salientar que a passagem do esquema de manutenção U para o esquema CR e posteriormente para o esquema SR originou:

- melhoria da utilização das centrais térmicas de base, por troca com as de ponta,
- melhoria da garantia do sistema electroprodutor através da redução da quantidade de energia não satisfeita pelo sistema,
- redução dos encargos com combustível e dos custos marginais de energia.

QUADRO I

Comparação de resultados obtidos para a exploração simulada do estádio 1986, na média dos 30 regimes hidrológicos disponíveis (1946-1975), para diferentes esquemas de manutenção.

Esq. de manut.	Produção Total (GWh)			Utilização da Pot. Inst. em Centrais Térm.(%)		E.N.S.(4)	L.O.L.P.(5)	C.M.E.(6)	enc. de comb.
utilizado	Para consum.	Hídrica	Térmica (1)	De Base (2)	De Ponta (3)	(GWh)	(%)	($/KWh)	($10^6$$)
U	21265	10173	11340	51.8	5.3	10.4	1.17	7.615	47264
CR	21265	10161	11344	52.5	3.7	3.3	0.56	6.765	46466
SR	21265	10178	11328	52.7	3.4	2.6	0.28	6.689	46383

(1) Inclui importação nocturna de energia.
(2) Engloba as centrais térmicas de Sines, Setúbal e Carregado.
(3) Engloba as centrais térmicas de Tunes e Alto de Mira.
(4) Energia não satisfeita (importação diurna de energia).
(5) "Loss of load probability".
(6) Custo marginal de energia.

QUADRO II

Estádio 1986
Esquema de Manutenção CR
Mapa de manutenção anual para a média dos 30 regimes disponíveis (1946-1975)

SEMANAS

```
                       1111111111222222222233333333334444444444555
Central Grupo 1234567890123456789012345678901234567890123456789012
SINES   1                11111111
SINES   2                        22222222
SETB    1               11111111
SETB    2                      22222222
SETB    3                              33333333
SETB    4                                      44444444
CARR    1      11111111
CARR    2        22222222
CARR    3                   33333333
CARR    4                                 44444444
CARR    5                                           55555555
CARR    6                                               66666666
TOUT    1     111111111111
TOUT    2                              222222222222
TOUT    3                                           333333333333
BHRR    1                                  1111111111
TGAT    1                                      111
TGAT    2                                         222
TGANT   1                                   111
TGANT   2                                                222
TGANT   3                                                   333
TGANT   4                                                      444
                       1111111111222222222233333333334444444444555
Central Grupo 1234567890123456789012345678901234567890123456789012
```

QUADRO III

Estádio 1986
Esquema de Manutenção SR

Mapa de manutenção anual para a média dos 30 regimes disponíveis (1946-1975)

SEMANAS

Central	Grupo	1111111111222222222233333333334444444444555 1234567890123456789012345678901234567890123456789012
SINES	1	11111111
SINES	2	22222222
SETB	1	11111111
SETB	2	22222222
SETB	3	33333333
SETB	4	44444444
CARR	1	11111111
CARR	2	22222222
CARR	3	33333333
CARR	4	44444444
CARR	5	55555555
CARR	6	66666666
TOUT	1	111111111111
TOUT	2	222222222222
TOUT	3	333333333333
BHRR	1	1111111111
TGAT	1	111
TGAT	2	222
TGANT	1	111
TGANT	2	222
TGANT	3	333
TGANT	4	444
Central	Grupo	1111111111222222222233333333334444444444555 1234567890123456789012345678901234567890123456789012

Bibliografia

A. Seca Teixeira, Victor Baptista, Carlos Meneses
"Modelos de Programação da Manutenção Preventiva de Grupos Térmicas em Estudos de Planeamento de Centros Produtores de Energia Eléctrica" - DT 91/81/OCPL - Maio de 1981

Investigação Operacional, Vol.7 nº1 (Junho/1987) 47 - 53

UM ALGORITMO DE SIMULAÇÃO-OPTIMIZAÇÃO PARA PROBLEMAS ESTOCÁSTICOS

J. Rodrigues Dias

Departamento de Matemática
Universidade de Évora
7000 Évora - Portugal

Resumo: Apresenta-se um algoritmo de simulação para optimizar uma função objectivo em que intervêm grandezas aleatórias. O algoritmo associa, por um lado, um método simples de pesquisa da solução e, por outro lado, um método de controlo estatístico da precisão dos resultados. Para tal, usando a técnica das variáveis antitéticas, fazem-se duas corridas paralelas a partir das quais se estima a precisão dos resultados obtidos e se estabelece o comprimento (que em princípio vai crescendo) das próprias corridas de simulação. Como critérios de paragem, consideram-se os erros relativos das variáveis independentes e o coeficiente de variação da função objectivo.

Como exemplo, apresenta-se o caso da minimização de um função no âmbito da fiabilidade e da inspecção de sistemas, que tem a particularidade de ter mais do que um mínimo relativo.

Abstract: We present an algorithm using simulation techniques to optimize an objective function depending on random variables. The algorithm combines a simple search method to obtain the solution and a statistical method to control the results precision, using the antithetic variables technique in two parallel runs. So, it is possible to estimate the run length in the next system simulation.

We present an example in which we minimize an objective function with two local minima, concerning a system inspection policy.

Keywords: Simulation, Optimization, Stochastic Problems, Statistical Techniques, Antithetic Variables.

1. Introdução

O problema da optimização é de grande importância em Investigação Operacional e, genericamente, em todas as ciências aplicadas. Por outro lado, o recurso crescente a técnicas de simulação, em especial se a complexidade das questões é grande, coloca a simulação em lugar cimeiro entre os métodos de análise e resolução de problemas (veja-se, por exemplo, [7]). Daí o interesse enorme do binómio simulação-optimização.

Como é sabido, a optimização de funções aparece quer em problemas de natureza determinística, quer em problemas de natureza probabilística. Nuns e noutros, a sua solução pode obter-se usando métodos determinísticos ou então usando técnicas probabilísticas. Para uma maior informação veja-se, por exemplo, [2], [3], [5], [6] e [14].

Contrariamente ao que se passa nos métodos determinísticos, parece-nos que na utilização de técnicas probabilísticas de simulação existe (ainda) uma certa "arte", para já não falar da discussão que elas podem suscitar quanto à precisão dos resultados obtidos e mesmo quanto à sua própria filosofia (veja-se [13]). É claro que isto em nada diminui a importância enorme da simulação nos dias de hoje, em especial tendo em conta as disponibilidades e as potencialidades crescentes dos meios informáticos.

O presente artigo, que resulta basicamente da experiência e das conclusões a que o autor foi chegando ao longo de trabalhos anteriores feitos na área (veja-se, por exemplo, [8], [9], [10] e [11]), pretende apresentar os aspectos fundamentais de um algoritmo de optimização em problemas em que há necessidade de simular grandezas aleatórias - problemas estocásticos. Assim, o algoritmo associa, por um lado, um método de pesquisa da solução e, por outro lado, um método de estimação da dispersão dos resultados, já que eles próprios são grandezas aleatórias. A conjugação dos dois métodos permite estabelecer para cada nova simulação o comprimento da respectiva corrida, de modo a que o valor final da função a optimizar se situe a um nível de precisão previamente estabelecido. Refira-se que esta é uma questão importante neste tipo de problemas (veja-se, por exemplo, [2]).

2. O Algoritmo de Simulação-Optimização

2.1. A função a optimizar

Consideremos um sistema em que actuam as variáveis aleatórias $\{X_i\}$ e as variáveis controláveis (independentes) $\{v_j\}$, para além de um conjunto de parâmetros não controláveis e eventualmente de um conjunto de restrições. Seja Y a resposta do sistema a analisar. Tem-se, então:

$$Y = Y(\{X_i\}, \{v_j\}), \quad i = 1, 2, 3,\ldots, k_1 \;\; ; \;\; j = 1, 2, 3,\ldots, k_2$$

ou seja, Y é uma função de $\{X_i\}$ e de $\{v_j\}$.

Considere-se, por outro lado, que estamos interessados em analisar a esperança matemática E(Y) de Y, pelo que se pode escrever:

$$E(Y) = F(\{v_j\})$$

em que E(Y), para além de depender naturalmente de $\{X_i\}$, depende dos valores que as variáveis controláveis v_j possam tomar. Para uma análise mais detalhada desta questão é conveniente ver, por exemplo, [4] e [5].

O nosso objectivo é, então, determinar um conjunto particular de valores $\{v_j^\circ\}$ de tal modo que E(Y) tome um valor óptimo.

A título de exemplo, que depois teremos a possibilidade de ver com mais pormenor, pode considerar-se o caso da minimização do custo total médio de funcionamento de um sistema por ciclo.

2.2. O método de pesquisa

Confrome se assinalou, o objectivo é então optimizar a função $E(Y) = F(\{v_j\})$, para o que é necessário determinar o conveniente conjunto de valores $\{v_j^\circ\}$. A partir daqui, sem perda de generalidades, admitir-se-á que se trata de um problema de minimização.

A ideia-base do método de pesquisa consiste em começar por fazer uma análise global da função no espaço das soluções que se consideram possíveis, a partir da qual se passa para uma análise local, cada vez mais restrita, nas zonas (pode ser apenas uma) onde a função tomar menores valores.

Por uma questão de simplicidade de exposição, considere-se o caso em que se tem apenas duas variáveis v_1 e v_2 e cujos valores optimizantes se encontram (ou se supõe que se encontram) nos intervalos (a_1, b_1) e (a_2, b_2), respectivamente. A partir daqui, define-se uma malha inicial cujos nós, em número a definir pelo utilizador, estão no que concerne a cada variável a uma mesma distância entre si. É no conjunto dos nós assim definidos que se vai fazer a análise global atrás referida, mediante o cálculo da função nesses pontos.

Conhecidos os valores da função F = E(Y) em cada um dos nós dessa malha inicial, vamos destes reter apenas aqueles em que a função tomar os menores valores, em número igualmente a definir pelo utilizador. Ora é justamente tomando estes nós para centros de novas malhas (submalhas) que se vai iniciar a análise local, calculando o valor da função nos respectivos nós (3^2 por cada submalha, no caso de 2 variáveis). Note-se que, relativamente à malha inicial, estes nós estão agora, no que diz respeito a cada uma das variáveis, a metade da distância a que estavam os nós iniciais. Assim, se d_1 e d_2 forem as distâncias entre nós consecutivos na malha inicial, em relação às variáveis v_1 e v_2, e se $d_1{}^1$ e $d_2{}^1$ forem as correspondentes distâncias na malha (submalha) seguinte, então tem-se: $d_1{}^1 = d_1/2$ e $d_2{}^1 = d_2/2$.

A partir da comparação dos valores que a função tomar em cada um dos nós das submalhas referidas, retêm-se apenas aqueles nós em que o valor da função for menor, a partir dos quais se definem analogamente novas submalhas que irão ser analisadas na iteração seguinte, etc. Este processo de análise local cada vez mais restrita acaba quando o erro relativo de cada variável (a definir adiante) for inferior a um valor previamente dado.

Generalizando, seja então:

d_j - distância, relativamente a cada uma das variáveis v_j, entre nós consecutivos da malha inicial;

$d_j{}^k$ - distância, relativamente a cada uma das variáveis v_j, entre nós consecutivos da malha (submalha) de ordem k (após a inicial), correspondente à iteração de ordem k;

$v_j{}^k$ - valor da variável v_j, na iteração de ordem k, no nó da malha onde a função tomar menor valor;

$e_j{}^k$ - erro relativo de v_j (estimativa) na iteração de ordem k, definido do seguinte modo:

$$e_j{}^k = d_j{}^k / v_j{}^k = d_j / (2^k \cdot v_j{}^k)$$

$e_j{}^*$ - erro relativo de v_j admissível, previamente definido.

Tendo em conta o que antes ficou dito, usa-se como critério de paragem na análise local a condição:

$$e_j{}^k < e_j{}^* \ , \ j = 1, 2, 3, \ldots, k_2$$

Refira-se, como curiosidade, que se 1000 é a distância entre nós consecutivos da malha inicial, no que diz respeito a uma dada variável, então ao fim de 10 iterações a distância entre nós consecutivos da respectiva submalha é inferior a 1:

$$d_j{}^{10} < d_j / 2^{10} = 1000 / 1024 < 1.$$

Este facto dá uma ideia da rapidez com que as zonas de análise local se vão reduzindo.

Note-se, já que esse resultado é importante na aplicação do algoritmo, que ao fim da iteração de ordem k é possível prever, àparte pequenas oscilações futuras nos valores de v_j, o número m de iterações que ainda é necessário realizar de tal modo que ao fim delas o erro relativo de cada uma das variáveis v_j seja inferior ao valor especificado. Assim, tem-se:

$$m = \max_j \ \mathrm{INT}\{\log[d_j{}^k / (e_j{}^* \, v_j{}^k)] / \log 2\} + 1$$

2.3. O método de controlo da precisão dos resultados

Recorde-se que o nosso objectivo é minimizar a esperança matemática da resposta Y do sistema que, sendo uma função de um conjunto $\{X_i\}$ de variáveis aleatórias, é ela própria uma grandeza aleatória.

Note-se que no caso em que não intervêm no sistema variáveis aleatórias, então Y é uma constante para cada conjunto particular de valores de $\{v_j\}$.

Seja, então, Z uma variável aleatória relativa ao funcionamento do sistema n vezes (por exemplo, durante n ciclos) definida do modo seguinte:

$$Z = \sum_{p=1}^{n} Y_p$$

em que Y_p, com $p = 1, 2, 3, \ldots, n$, são variáveis aleatórias (que supomos independentes) com a mesma distribuição de Y.

Sejam ainda Z_1 e Z_2 duas variáveis aleatórias com a mesma distribuição de Z. Tem-se então que

$$\bar{Y} = (Z_1 + Z_2) / (2n)$$

é um estimador centrado de $F = E(Y)$ pelo que, se z_1 e z_2 forem dois valores observados de Z, então $\bar{y} = (z_1 + z_2) / (2n)$ é uma estimativa centrada da função a minimizar F.

Por outro lado, representando por V(X) a variância da variável aleatória genérica X, tem-se:

$$\begin{aligned} V(Y) &= V(Z_1 + Z_2) / (4n^2) \\ &= \{ V(Z_1) + V(Z_2) + 2\Upsilon [V(Z_1) . V(Z_2)]^{1/2} \} / (4n^2) \\ &= V(Z) . (1 + \Upsilon) / (2n^2) \end{aligned}$$

em que Υ é o coeficiente de correlação de Z_1 e Z_2.

No caso em que Z_1 e Z_2 estiverem negativamente correlacionadas ($\Upsilon < 0$), então tem-se:

$$V(\bar{Y}) < V(Z) / (2n^2)$$

O coeficiente de variação de $\bar{Y}$ é dado por:

$$C_v (\bar{Y}) = [V(\bar{Y})]^{1/2} / E(\bar{Y})$$

pelo que, sendo $\Upsilon < 0$, se tem:

$$C_v (\bar{Y}) < [V(Y) / (2n)]^{1/2} / E(\bar{Y})$$

Querendo obter um coeficiente de variação de $\bar{Y}$ inferior a um dado valor $C_v{}^*$ previamente definido, então deverá tomar-se:

$$n = \mathrm{INT} \{ [V(Y) / 2] / [E(\bar{Y}) C_v{}^*] \}^2 + 1$$

É claro que para calcular este valor de n, que define o comprimento da corrida de simulação, deverá conhecer-se o valor de V(Y), o que na prática não acontece, naturalmente. Pode, no entanto, calcular-se uma estimativa. De facto, sendo V(Y) = V(Z) / n, e podendo considerar-se Z normal ou aproximadamente normal (pelo teorema do limite central), pode-se escrever:

$$V(Z) = [E(R_2) / d_2]^2$$
$$V(Y) = [E(R_2) / d_2]^2 / n$$

em que $E(R_2)$ é a esperança matemática da amplitude R_2 de uma amostra de tamanho 2 e d_2=1.128 é um factor apropriado (veja-se, por exemplo, [1]). Mas, como é evidente, a utilização das igualdades anteriores exige agora o conhecimento de uma estimativa de $E(R_2)$. Neste caso, por simplicidade e não só, como se verá a seguir, vai-se adoptar o valor da amplitude de uma amostra (Z_1, Z_2) de Z.

Recorde-se que o objectivo principal neste ponto do nosso trabalho é determinar uma estimativa do valor da função F através de $\bar{Y}$, pelo que é conveniente que este estimador tenha uma variância o menor possível. Ora já se viu que se Z_1 e Z_2 estiverem negativamente correlacionadas, então verifica-se:

$$\begin{aligned} V(\bar{Y}) &< V(Z) / (2n^2) \\ &< V(Y) / (2n) \end{aligned}$$

Daqui conclui-se que em termos de aplicação é preferível utilizar duas corridas paralelas de simulação negativamente correlacionadas, cada uma delas de comprimento n, do que uma só corrida de comprimento 2n, já que nesse caso se verifica uma redução de variância. Esta correlação negativa entre Z_1 e Z_2 poderá conseguir-se utilizando a técnica das variáveis antitéticas (veja-se, por exemplo, [15]). Refira-se que, de entre as técnicas de redução de variância que o autor já usou, esta foi aquela que melhores resultados produziu.

Em conclusão, pode então dizer-se que a utilização de duas corridas paralelas de simulação negativamente correlacionadas permite atingir dois objectivos importantes: por um lado, reduzir a variância do estimador da função a minimizar e, por outro lado, calcular de uma forma muito simples uma estimativa (grosseira, por excesso) do respectivo coeficiente de variação, a partir da qual é possível calcular o comprimento dessas mesmas corridas de simulação.

2.4. Interligação entre o método de pesquisa e o método de controlo da precisão dos resultados

Já se viu anteriormente que o método de pesquisa da solução se considera o critério de paragem:

$$e_j^k = d_j / (2^k v_j^k) < e_j^*$$

Por outro lado, tendo em conta o carácter aleatório da resposta do sistema, neste algoritmo vai-se adoptar um outro critério de paragem: o coeficiente de variação do estimador $\bar{Y}$ de E(Y) deve ser inferior a um dado valor previamente especificado:

$$C_v(\bar{Y}) < C_v^*$$

Deste modo, o algoritmo de simulação-optimização termina quando se verificarem simultaneamente os dois critério de paragem.

Ora, tendo em conta que a adopção do primeiro critério permite determinar no fim de cada iteração o número previsível de iterações m ainda necessárias, e que, por outro lado, o método de controlo da precisão dos resultados utilizado permite determinar o comprimento n da corrida de simulação necessário para atingir um dado coeficiente de variação, então é possível adoptar a seguinte metodologia na implementação do algoritmo:

a) Determinação, no fim da iteração de ordem k, do número previsível m de iterações ainda necessárias de tal modo que $e_j^{k+m} < e_j^*$, com m dado pela fórmula que foi anteriormente apresentada.

b) Determinação, no fim da iteração de ordem k, de uma estimativa do coeficiente de variação C_v^k.

c) Tendo presente que o objectivo é obter na ietração de ordem k+m um coeficiente de variação inferior a C_v^*, e supondo (para efeitos de cálculo) uma variação linear a partir da iteração de ordem k no coeficiente de variação, então o valor deste na iteração de ordem k+1 deverá ser:

$$C_v^{k+1} = C_v^k - (C_v^k - C_v^*) / m$$

d) Calculado o valor desejável do coeficiente de variação na iteração de ordem k+1, é então possível determinar o comprimento n da respectiva corrida de simulação.

Refira-se que esta sequência de operações se repete no fim de cada iteração, pelo que a determinação de um novo valor de n tem sempre em conta os últimos resultados obtidos, que em princípio apresentam em relação aos anteriores uma maior precisão.

Note-se ainda que esta determinação automática de n só não é feita, naturalmente, na primeira corrida de simulação, caso em que o utilizador tem que indicar um valor (que, conforme se verá no exemplo a seguir, não é determinante em termos de resultados finais, dado o ajuste automático que depois vai sendo feito).

3. Um Exemplo de Aplicação

Consideremos o problema que em termos esquemáticos pode ser apresentado do modo seguinte:

a) Seja um sistema com apenas dois estados - um de bom e outro de mau funcionamento;

b) Este último, resultante de uma avaria, é apenas conhecido fazendo inspecções, que vamos supor de duração nula e periódicas;

c) A duração do estado de bom funcionamento (tempo de vida do sistema) é uma variável aleatória, cuja função de distribuição F(t) é conhecida;

d) Existe um custo C_1 por cada inspecção e um custo C_2 por cada unidade de tempo de mau funcionamento não detectado;

e) Finalmente, considera-se que a probabilidade de erro em cada inspecção é nula.

A questão que se põe então é determinar o período P entre inspecções de tal modo que o custo total médio por ciclo, que vamos designar por E(C), resultante do custo das inspecções e do custo de mau funcionamento, seja mínimo. Considera-se que um ciclo começa com o sistema em estado de novo e acaba quando a avaria é detectada. Para mais pormenores, veja-se [12].

Mostra-se que o valor de P que minimiza E(C) satisfaz a equação:

$$\sum_{k=0}^{\infty} R(kP) - (r + P) \sum_{k=1}^{\infty} kf(kP) = 0 \quad , \qquad r = C_1/C_2$$

em que $R(t) = 1\text{-}F(t)$ é a função de fiabilidade do sistema e $f(t)$ é a função densidade de probabilidade.

Esta equação pode admitir mais do que uma solução, como é o caso em que se considera a distribuição de Weibull, com $R(T) = [\exp(t / \alpha)]^{\beta}$, $\beta = 5$, $E(T) = 1$, e sendo $C_1 = .2$ e $C_2 = 1$. Nestas condições, a função a minimizar E(C) apresenta dois mínimos com valores muito aproximados, um para P≈.6 e outro para P≈ 1.3.

Embora em [12] não se tenha aoptado este processo, vamos aqui resolver este problema de simulação-optimização utilizando o algoritmo descrito. Para esse efeito, considerou-se:

a) $e_1^* = 1\%$;
b) $C_v^* = 2\%$;
c) Número inicial de pontos nos diferentes intervalos (a_1, b_1): 15 ;
d) Número de submalhas para análise em cada iteração: 2.

Para efeitos comparativos, consideraram-şe diferentes intervalos (a_1, b_1) e diferentes valores de n_0 para o comprimento inicial das corridas de simulação. Os resultados obtidos apresentam-se em Quadro.

Refira-se que o valor mínimo da função F = E(C) é .6311 e é obtida para P = 1.3246.

(a_1, b_1)	n_0: 50	100	200	400
(.5, 1.5)	P = 1.3303 F = .6147	P = 1.3303 F = .6034	P = 1.3303 F = .6352	P = 1.3303 F = .6353
(.5, 1.6)	P = 1.3249 F = .6269	P = 1.3200 F = .6364	P = 1.3249 F = .6112	P = 1.3151 F = .6352
(.5, 1.5)	P = 1.3241 F = .5939	P = 1.3124 F = .6314	P = 1.3357 F = .6360	P = 1.3124 F = .6363

Quadro: Valores obtidos para o custo total médio mínimo F = E(C) e para o correspondente período de inspecção P, sendo os valores exactos: $F_{min} = .6311$; $P_{opt} = 1.3246$.

Conforme se verifica, apesar de em qualquer dos intervalos utilizados a função F = E(C) apresentar dois mínimos, a verdade é que o algoritmo conduziu sempre ao mínimo absoluto (valor aproximado). A esta mesma situação poderia não se chegar se se utilizasse um outro algoritmo. Por outro lado, se é verdade que a escolha de um valor elevado de n_0 na primeira corrida de simulação garante em princípio melhores resultados, também é verdade que o ajuste automático dos valores de n nas corridas seguintes acaba por tornar pouco relevante esse mesmo valor inicial n_0.

4. Considerações Finais

Antes de mais, deve acentuar-se que este algoritmo é de implementação muito simples, podendo correr com facilidade em qualquer microcomputador, até pelo facto de exigir uma memória extremamente reduzida. Este facto é tanto mais importante quanto é certo que a microinformática está hoje ao alcance de todos.

No entanto, convém salientar que o método de pesquisa da solução pode exigir um elevado tempo de computação, como consequência do número eventualmente grande de pontos (nós das malhas) em que a função tem que ser calculada. De facto, na análise global o número de nós é N^{k_2}, em que N é o número de pontos que se considera no intervalo (a_j, b_j) e k_2 é o número de variáveis. Na análise local, o número de nos é $S.3^{k_2}$, em que S é o número de submalhas que se consideram.

Parece preferível, em especial se o número de variáveis é grande, utilizar na análise global uma malha bastante larga (número reduzido de pontos em cada intervalo (a_j, b_j)) e aumentar o número de submalhas nas iterações seguintes. De resto, note-se que até o facto de considerar mais que uma submalha tem vantagens imediatas: por um lado, reduz a margem de erro na pesquisa da solução e, por outro lado, dá a possibilidade, quando é caso disso, de se obter mais do que um mínimo da função. Este aspecto parece-nos importante em termos de comparação de resultados, até numa perspectiva de análise de sensibilidade.

É também de acentuar que o método da bissecção utilizado na construção das submalhas permite calcular o número previsível de iterações ainda necessárias, o que dá a possibilidade de controlar automaticamente o comprimento da corrida de simulação até se atingir no final a precisão previamente estabelecida. Este é talvez o aspecto que nos parece mais importante, já que permite trabalhar com menor precisão nas iterações iniciais (basta considerar um valor reduzido no comprimento da corrida inicial), o que tem como consequência um menor tempo de computação. A este processo de ajuste automático do comprimento das corridas talvez se pudesse chamar "método da imagem variável com precisão ajustada" em oposição ao "método da imagem fixa" (veja-se [7]).

Refira-se ainda como aspecto positivo que o método de pesquisa pode provocar a saída de um intervalo (a_j, b_j) inicialmente dado, desde que a solução nele não esteja contida, o que pelo menos funciona como um alerta.

Note-se, por fim, que o algoritmo poderá ser aplicado, com as cautelas devidas, no caso em que existam restrições, bastando para isso considerar as convenientes funções penalidade. Do mesmo modo, e como caso particular que é, poderá ser também aplicado à optimização de problemas determinísticos.

Referências

[1] Grant, E.L. & Leavenworth, R.S. (1980) • *Statistical Quality Control*, Mc Graw-Hill (5ª edição).

[2] Guimarães, R.C. (1983) • Simulação-Optimização, *Investigação Operacional*, vol.3, nº1, pp.29-52.

[3] Himmelblau, D.M. (1972) • *Applied Nonlinear Programming*, Mc Graw-Hill.

[4] Kleijnen, J.P.C. (1975) • *Statistical Techniques in Simulation*, vol.II, Marcel Dekker.

[5] Mihram, G.A. (1972) • *Simulation: Statistical Foundations and Methodology*, Academic Press.

[6] Moitinho, J.P. (1982) • Método de Optimização por Simulação, *Investigação Operacional*, vol.2, nº1, pp.43-48.

[7] Oliveira, R. (1982) • Inquérito sobre Aplicações de I.O. em Portugal, *Investigação Operacional*, vol.2, nº2, pp.11-12.

[8] Rodrigues Dias, J. (1979) • A Simulação na Gestão de Stocks, *Economia e Sociologia*, nº27, pp.59-73.

[9] Rodrigues Dias, J. (1979) • *Aplicação do Método de Monte-Carlo ao Cálculo dos Valores de π e de 'e'*, Publicações Universidade de Évora, Ciâncias Exactas, nº1.

[10] Rodrigues Dias, J. (1980) • Método de Simulação Aplicado à Minimização de Custos na Substituição de Lâmpadas, *Electricidade*, nº150, pp.158-167.

[11] Rodrigues Dias, J. (1981) • Optimização de Intervalo de Amostragem em Controlo Estatístico de Qualidade, *Ordem dos Engenheiros, Congresso 81*, Tema 1, Comunicação16.

[12] Rodrigues Dias, J. (1983) • Influence de la Période d'Inspection sur les Coûts dans l'Inspection Périodique de Systéms, *Revue de Statistique Appliquée*, vol.XXXI, nº4, pp.5-15.

[13] Saliby, E. (1982) • Uma Revisão de Fundamentos da Simulação: O Uso Incorrecto de Amostragem Aleatória Simples, *Pesquisa Operacional*, vol.2, nº2, pp.1-16.

[14] Shimizu, T. (1985) • *Simulação em Computador Digital*, Edgard Blucher.

[15] Tocher, K.D. (1963) • *The Art of Simulation*, Hodder and Stoughton.

Investigação Operacional, Vol.7 nº1 (Junho/1987) 55 - 60

UMA NOTA SOBRE A FUNÇÃO DE PRODUÇÃO DE LEONTIEF

Fernando de Jesus

Instituto Superior de Economia
Universidade Técnica de Lisboa

Resumo: Dado que a função de produção de LEONTIEF se reveste de inegável importância teórica e prática quer ao nível microeconómico quer ao nível macroeconómico, sendo não raras vezes insuficientemente tratada nos manuais de teoria económica, pretende-se com esta nota referir os aspectos matemáticos básicos que se revelam de maior interesse para o economista.

Abstract: Since the Leontief production function has an extraordinary importance either theoretical or practical both at the microeconomic and macroeconomic levels, being not seldom insufficiently dealt in the textbooks of economic theory, the intention of this note is to refer the basic mathematical aspects which are considered to be the most interesting for the economist.

Keywords: Funções de produção (production functions), função de produção de Leontief (the Leontief production function), teoria da produção (the theory of production), teoria da empresa (theory of the firm).

Suponhamos que a quantidade q de um produto, que pode ser obtido a partir de uma combinação de n factores produtivos, satisfaz as relações tecnológicas

$$x_j = a_j\ q\ (\ j = 1,\dots, n\ ),$$

onde a_j (coeficiente técnico) designa uma constante positiva e x_j denota a quantidade utilizada do j-ésimo factor.

É óbvio que, para qualquer combinação produtiva $(x_1,\dots, x_n) \in R^n_+$, a quantidade máxima de produto que se pode obter (função de produção) é dada por

$$q \ = \ f\ (x_1,\dots, x_n) = \min\ (x_1/a_1,\dots, x_n/a_n).$$

Esta função de produção, que tem sido designada na literatura económica por função de produção com coeficientes técnicos fixos, ou função de produção de LEONTIEF, embora contínua, não possui as propriedades de diferenciabilidade admitidas na teoria neoclássica da empresa.

Dado que a função de produção de LEONTIEF se reveste de inegável importância teórica e prática quer ao nível microeconómico quer ao nível macroeconómico, sendo não raras vezes insuficientemente tratada nos manuais de teoria económica, pretende-se com esta nota referir os aspectos matemáticos básicos que se revelam de maior interesse para o economista.

Considerando apenas, por razões de simplicidade, dois facores produtivos, a função de produção

$$q = f(x_1, x_2) = \min(x_1/a_1, x_2/a_2)$$

pode também apresentar-se na forma

$$q = f(x_1, x_2) = \begin{cases} x_1/a_1 & (x_2 \geq \frac{a_2}{a_1} x_1) \\ x_2/a_2 & (x_2 < \frac{a_2}{a_1} x_1) \end{cases}$$

É fácil justificar que a função é côncava e, portanto, apresenta isoquantas convexas cuja representação se faz na Fig. 1.

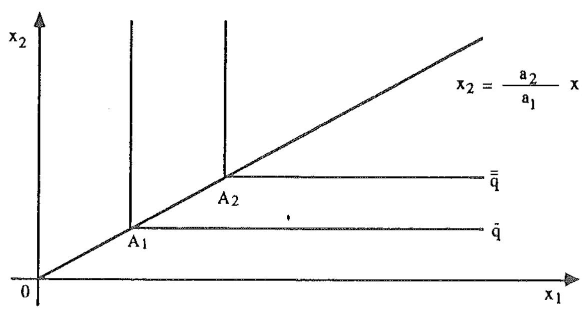


Fig. 1

As isoquantas definem ângulos rectos de lados paralelos aos eixos coordenaos, com vértices situados sobre a recta $x_2 = \frac{a_2}{a_1} x_1$.

É evidente que as combinações produtivas tecnicamente eficazes, i.e., as que permitem obter a quantidade q do produto sem desperdício de qualquer dos facotres, encontram-se sobre a recta $x_2 = \frac{a_2}{a_1} x_1$.

Esta representa, simultaneamente, as linhas de cume e o caminho de expansão a longo prazo. Os factores produtivos são complementares, combinando-se optimamente na razão fixa $x_2/x_1 = a_2/a_1$.

Também é de fácil verificação a homogeneidade linear da função, o que significa que a tecnologia traduzida pela função apresenta rendimentos de escala constantes.

Sendo $f_1(x_1, x_2)$ e $f_2(x_1, x_2)$ as derivadas parciais (produtividades marginais) de $f(x_1, x_2)$, tem-se

$$f_1(x_1, x_2) = \begin{cases} 1/a_1 & (x_2 > \frac{a_2}{a_1} x_1) \\ 0 & (x_2 < \frac{a_2}{a_1} x_1) \end{cases}$$

$$f_2(x_1, x_2) = \begin{cases} 0 & (x_2 > \frac{a_2}{a_1} x_1) \\ 1/a_2 & (x_2 < \frac{a_2}{a_1} x_1) \end{cases}$$

A função não possui derivadas parciais sobre a recta $x_2 = \frac{a_2}{a_1} x_1$, o que desrespeita uma das hipóteses básicas da teoria neoclássica.

É interessante apresentar as imagens das produtividades total, média e marginal dos factores. Considerando apenas o primeiro factor e fixando o segundo no nível $\bar{x}_2$, a produtividade total do factor 1 fica definida por

$$f(x_1, \bar{x}_2) = \begin{cases} x_1/a_1 & \left(x_1 \leq \frac{a_1}{a_2} \bar{x}_2 \right) \\ \bar{x}_2/a_2 & \left(x_1 > \frac{a_1}{a_2} \bar{x}_2 \right) \end{cases}$$

cuja imagem se apresenta na Fig. 2.

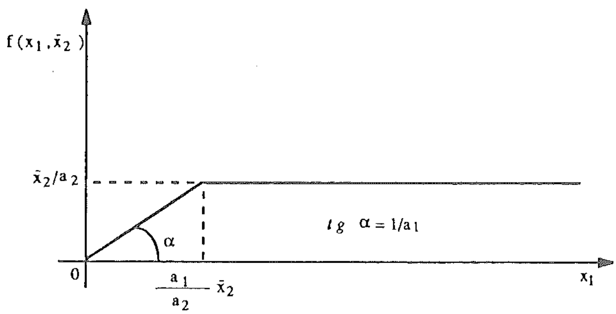


Fig. 2

A produtividade média $f(x_1, \bar{x}_2)/x_1$ vem definida por

$$f(x_1, \bar{x}_2)/x_1 = \begin{cases} 1/a_1 & \left(x_1 \leq \frac{a_1}{a_2} \bar{x}_2 \right) \\ \bar{x}_2/a_2\, x_1 & \left(x_1 > \frac{a_1}{a_2} \bar{x}_2 \right) \end{cases}$$

e representa-se na Fig. 3.

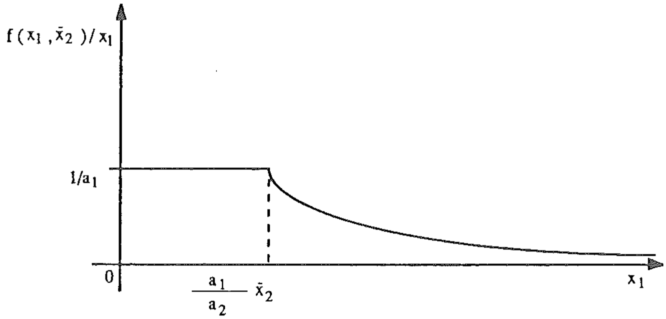


Fig. 3

Finalmente, a produtividade marginal do factor 1 é

$$f_1(x_1, \bar{x}_2) = \begin{cases} 1/a_1 & \left(x_1 < \frac{a_1}{a_2} \bar{x}_2 \right) \\ 0 & \left(x_1 > \frac{a_1}{a_2} \bar{x}_2 \right) \end{cases}$$

e representa-se na Fig. 4.

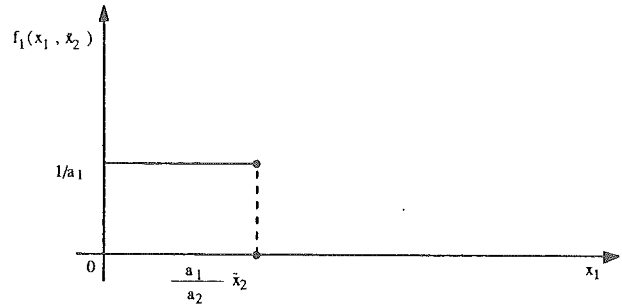


Fig. 4

As elasticidades da produção são dadas por

$$\frac{Eq}{Ex_1} = \begin{cases} 1 & (x_2 > \frac{a_2}{a_1} x_1) \\ 0 & (x_2 < \frac{a_2}{a_1} x_1) \end{cases} \qquad \frac{Eq}{Ex_2} = \begin{cases} 0 & (x_2 > \frac{a_2}{a_1} x_1) \\ 1 & (x_2 < \frac{a_2}{a_1} x_1) \end{cases}$$

e a elasticidade de escala vem dada por

$$\epsilon = \frac{Eq}{Ex_1} + \frac{Eq}{Ex_2} = 1,$$

para $x_2 \neq \frac{a_2}{a_1} x_1$.

A relação marginal de substituição do factor 2 pelo factor 1 é definida por

$$R_1^2 = -\frac{dx_2}{dx_1} = f_1/f_2 = 0 \qquad (x_2 < \frac{a_2}{a_1} x_1)$$

e a relação marginal de substituição do factor 1 pelo factor 2 é definida por

$$R_2^1 = -\frac{dx_1}{dx_2} = f_2/f_1 = 0 \qquad (x_2 > \frac{a_2}{a_1} x_1).$$

Embora não seja possível, em rigor, utilizar a definição de elasticidade de substituição, é costume afirmar que $\sigma_{12} = \sigma_{21} = 0$, atendendo a que a função de produção de LEONTIEF pode considerar-se como limite da função de produção *ESC* homogénea de grau um,

$$q = \gamma \left[(1 - \delta) x_1^{-\alpha} + \delta x_2^{-\alpha} \right]^{-1/\alpha},$$

quando $\alpha \to +\infty$, mediante a escolha de convenientes unidades de medida para os factores produtivos.

Sendo w_1 e w_2 os preços unitários dos factores produtivos, consideremos o problema fundamental da teoria da empresa

$$\min \quad c = w_1 x_1 + w_2 x_2$$

com

$$f(x_1, x_2) = q.$$

A solução óptima é obviamente

$$x_1^* = a_1 q$$
$$x_2^* = a_2 q$$

e, portanto, a função custo total a longo prazo associada à função de produção de LEONTIEF é

$$c^*(q, w_1, w_2) = (a_1 w_1 + a_2 w_2)\, q,$$

o que implica que os custos médio e marginal a longo prazo são constantes e iguais a $a_1 w_1 + a_2 w_2$.

A determinação da função custo total a curto prazo também não oferece dificuldades.

Fixando, por exemplo, o factor 2 no nível $\bar{x}_2$, teremos de resolver o problema

$$\min \quad w_1x_1 + w_2\bar{x}_2$$

com

$$f(x_1\bar{x}_2) = q \qquad (0 \le q \le \bar{x}_2/a_2).$$

Da última equação vem

$$x_1 = a_1 \, q$$

e a função-custo total a curto prazo vem definida por

$$c^* (q, w_1, w_2, \bar{x}_2) = a_1w_1 \, q + w_2\bar{x}_2 ,$$

para $(0 \le q \le \bar{x}_2/a_2)$. Do custo total a curto prazo obtêm-se facilmente os correspondentes custo médio e marginal:

$$\frac{c^* (q, w_1, w_2, \bar{x}_2)}{q} = a_2w_1 + \frac{w_2 \bar{x}_2}{q}$$

$$\frac{\partial c^* (q, w_1, w_2, \bar{x}_2)}{\partial q} = a_1w_1 ,$$

funções também definidas no intervalo $[0, \bar{x}_2/a_2]$.

Como se sabe, a linearidade em q da função-custo total implica dificuldades na teoria neoclássica da empresa, designadamente no problema da determinação da produção que maximiza o lucro em regime de concorrência perfeita. De facto, no horizonte de longo prazo, notando que o lucro é definido por

$$\pi = pq - c^* (q, w_1, w_2) = (p - a_1w_1 - a_2w_2)\, q,$$

onde p denota o preço unitário do produto, verifica-se facilmente o seguinte:

a) se $p > a_1w_1 + a_2w_2$, a empresa terá tendência para expandir indefinidamente a produção, em consequência do lucro crescer incessantemente com a quantidade produzida;

b) se $p < a_1w_1 + a_2w_2$, o máximo do lucro corresponderá a $q = 0$, i.e., a empresa nada produzirá.

c) se $p = a_1w_1 + a_2w_2$, única hipótese compatível com a condição de primeira ordem

$$\frac{\partial \pi}{\partial q} = p - \frac{\partial c^*}{\partial q} = 0,$$

o lucro é nulo para qualquer nível da produção, i.e., não há uma solução única para o problema de maximização do lucro.

A maximização da produção para um custo fixado pode formalizar-se do modo seguinte:

$$\max \quad q = f(x_1, x_2),$$

com

$$w_1x_1 + w_2x_2 = c.$$

O problema possui solução simples. De facto, como a solução óptima tem de existir sobre a recta $x_2 = (a_2/a_1)\, x_1$, a resolução do sistema

$$\begin{cases} x_2 = \dfrac{a_2}{a_1} x_1 \\ w_1x_1 + w_2x_2 = c \end{cases}$$

conduz à solução

$$x_1^* = \frac{a_1 \, c}{a_1w_1 + a_2w_2} , \qquad x_2^* = \frac{a_2 \, c}{a_1w_1 + a_2w_2} .$$

A produção óptima é então dada por

$$q^* = x_1^*/a_1 = x_2^*/a_2 = \frac{c}{a_1w_1 + a_2w_2} .$$

A função de produção de LEONTIEF pode ser facilmente generalizada, considerando a tecnologia definida por

$$x_j = f_j(q) \qquad (j = ,\dots, n),$$

onde as funções $f_j(q)$ são contínuas e deriváveis, satisfazendo também às seguintes condições:

$$f_j(q) \ge 0, \;\; f_j(0) = 0, \;\; f'_j(q) \ge 0$$

A monotonia das funções $f_j(q)$ implicará a existência das funções inversas

$$q = f_j^{-1}(x_j) \qquad (j = 1,\dots, n)$$

e a função de produção pode ser dada por

$$q = \min\left[\, f_1^{-1}(x_j),\dots, f_n^{-1}(x_n) \,\right].$$

Voltando a considerar apenas dois factores, e fazendo por comodidade $\phi_1(x_1) = f_1^{-1}(x_1)$ e $\phi_2(x_2) = f_1^{-1}(x_2)$, a função de produção

$$q = \min\left[\, \phi_1(x_1), \phi_2(x_2) \,\right]$$

pode apresentar-se na forma

$$q = \begin{cases} \phi_1(x_1) & (\phi_1(x_1) \leq \phi_2(x_2)) \\ \phi_2(x_2) & (\phi_1(x_1) > \phi_2(x_2)) \end{cases}$$

e as suas isoquantas definem ainda ângulos rectos de lados paralelos aos eixos coordenados, com vértices situados sobre a curva $\phi_1(x_1) = \phi_2(x_2)$ (Fig. 5).

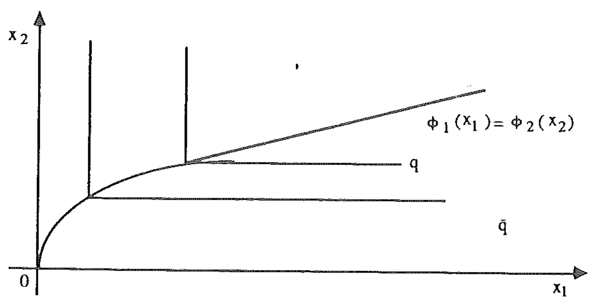


Fig. 5

Um caso interessante é aquele em que a tecnologia é definida pelas relações

$$x_j = a_j \quad q^{1/h} \qquad (j = 1,\dots, n),$$

com $h > 0$. Trata-se de uma tecnologia que apresenta rendimentos de escala crescentes, decrescentes e constantes, consoante $h > 1$, $h < 1$ e $h = 1$. Neste último caso, é a tecnologia de LEONTIEF.

A função de produção é definida por

$$q = \min\left[\, (x_1/a_1)^h,\dots, (x_n/a_n)^h \,\right]$$

No caso de dois factores, a função pode escrever-se na forma

$$q = \begin{cases} (x_1/a_1)^h & \left(x_2 \geq \dfrac{a_2}{a_1} x_1\right) \\ (x_2/a_2)^h & \left(x_2 < \dfrac{a_2}{a_1} x_1\right) \end{cases}$$

e as isoquantas, tal como na Fig. 1, definem ângulos rectos de lados paralelos aos eixos coordenados, com vértices situados sobre a recta $x_2 = \dfrac{a_2}{a_1} x_1$.

Investigação Operacional, Vol.7 nº1 (Junho/1987) 61 - 65

REDES DE SERVIÇOS PÚBLICOS E COBERTURA DO TERITÓRIO

José Manuel Viegas

CESUR - Instituto Superior Técnico

1 - Conceitos Fundamentais

A existência de serviços públicos nas sociedades, e particularmente nas sociedades modernas, está associada ao reconhecimento da existência de necessidades colectivas e de necessidades individuais que devem ser satisfeitas a todos os cidadãos.

As necessidades colectivas são satisfeitas através da produção de bens ou serviços públicos que têm como características fundamentais a impossibilidade de apreciação total ou parcial do bem ou serviço pelo beneficiário e, como consequência, a não rivalidade entre beneficiários.

Ou seja, um serviço público (neste sentido) é tal que, ao ser produzido, torna-se disponível - por vezes obrigatório e inescapável - para todos os potenciais beneficiários que por ele sejam abrangidos. Os casos mais típicos são a Defesa Nacional, o policiamento, a (rádio) difusão de informação por exemplo.

O que acontece é que, por virtude da não apropriação do serviço pelo beneficiário, este não é indentificável e portanto não é possível descriminar (seleccionar) os beneficiários.

As únicas formas de discriminação possíveis neste caso dizem respeito à cobertura ou não de um determinado território, durante um dado período de tempo, e ainda ao nível (intensidade e qualidade) dessa cobertura. Mas, repete-se, a discriminação não é feita a nível do indivíduo e sim a nível do território. Por isso, não pode ser cobrado directamente qualquer preço pelo usufruto desse bem ou serviço.

No outro extremo do espectro estão chamados bens ou serviços privados, destinados a satisfazer necessidades individuais, e em que a apropriação do bem pelo beneficiário é total, podendo portanto a rivalidade entre esses consumidores existir e ser forte. A possibilidade de identificar os consumidores leva naturalmente à sua discriminação, geralmente através da disposição ou não de pagar um preço por esse bem ou serviço, mas não necessariamente. Outros tipos de discriminação frequentes são a pertença a grupos fechados ou abertos (sócio ou membro de uma colectividade, idade, religião, etc).

Existe um terceiro grupo, a que é habitual chamar bens ou serviços semi-públicos, em que se incluem todos aqueles que satisfazem necessidades que, para além de serem sentidas pelo indivíduo, são também consideradas pela sociedade (ou pelos seus dirigentes pelo menos) como devendo ser satisfeitas. Este grupo abrange um leque considerável de serviços para os quais o Estado garante, se não a sua produção, pelo menos a sua provisão, isto é, o Estado ou produz directamente ou tutela e regulamenta essa produção.

O número destes serviços tem vindo a crescer à medida que a sociedade vai tendo uma organização mais complexa, precisamente na medida em que o aumento das inter-relações e o desejo de manter ou aumentar a eficácia global de funcionamento da sociedade obrigam a uma maior uniformização (standardização) e a uma maior eficácia de funcionamento de cada um dos indivíduos.

Se quisermos citar exemplos bastará pensar nas áreas da educação, da saúde, da cultura, e, de um modo geral, todas as infraestruturas e os transportes públicos.

A referida complexidade tem vindo a provocar o aumento seja do nível (qualidade e sofisticação) dos serviços prestados como também da intromissão do Estado como regulamentador da prestação de serviços em novas áreas.

A própria natureza mista destes serviços, se bem que permita a identificação dos consumidores e dê lugar à existência de competitividade entre eles, tem um modo geral conduzido à prática de preços "sociais", inferiores muitas vezes ao custo de produção, de modo a permitir o acesso de toda a população a tal bem ou serviço.

Surge assim um problema - fixação de preços - que deve ser resolvido tendo em vista a satisfação de dois objectivos contraditórios: a cobertura tão grande quanto possível dos custos de produção do serviço pelas suas receitas, e a cobertura da maior parte possível da população pelo serviço (em termos do acesso efectivo a tal serviço).

Se o não cumprimento integral do primeiro objectivo implica o recurso a financiamentos do Estado (e portanto dos contribuintes em geral, consumidores ou não do serviço em causa), o não cumprimento integral do segundo implicará a existência de contribuintes que estão a ser descriminados por não terem acesso a um serviço que a comunidade considera corresponder à satisfação de necessidades fundamentais, e portanto subsidia. Uma situação semelhante, que existe frequentemente, corresponde à existência de acesso ao serviço mas a níveis de tal maneira degradados de qualidade do serviço que não pode deixar de considerar-se a existência de discriminação.

Embora o primeiro destes dois tipos de problemas (os "prejuízos" das empresas públicas) seja o mais frequentemente discutido, propomo-nos aqui tratar o segundo (níveis de cobertura do território/populações pelos serviços semi-públicos e públicos).

2 - Qualidade dos Serviços e Nível de Cobertura

É naturalmente difícil de medir a qualidade (e a sua variação) de um serviço público ou semi-público que o Estado presta à população nas várias zonas do território. Essa dificuldade tem a ver com o facto de frequentemente o produto fornecido não ser facilmente quantificável, como é o caso dos serviços de cultura, educação e saúde.

No entanto, e para efeitos de detecção de situações de discriminação, será possível substituir a produção do produto pela medição ou dos seus factores de produção ou de alguns parâmetros (quantitativos) caracterizadores desse mesmo produto.

Algumas dessas medidas podem ser por exemplo número de funcionários por utente, taxas médias de sucesso (em escolas ou em equipamentos de saúde), ou outras.

Não sendo uma informação perfeita, será no entanto, na maioria dos casos, suficiente para detectar boa parte das situações de discriminação.

A recolha e análise comparativa destes indicadores é da maior utilidade já que torna possível a detecção de situações em que os utentes de determinadas instalações (e independente da distância que têm de percotrrer até aí) são sistematicamente pior servidos que os que a outras recorrem.

Se os aspectos de vantagem e desvantagem resultantes das distâncias entre os utentes e os equipamentos são de mais difícil ultrapassagem, estes outros são pelo contrário provenientes ou de sub-dotação de recursos humanos e/ou de capital ou de má gestão da instalação, sendo portanto mais fácil através de uma política global para a rede de equipamentos e sem recurso a medidas capital-intensivas, atenuar as diferenças entre os níveis (qualidade) de serviço prestados em cada uma das instalações.

Mas é preciso não esquecer que, visto do lado do consumidor, há uma outra característica fundamental na percepção do nível de serviço que lhe é fornecido, nomeadamente a distância (ou tempo) que tem de percorrer até ter acesso a esse serviço.

De facto a distância funciona como restritor maior do acesso aos equipamentos e serviços públicos, podendo dizer-se que para qualquer tipo de serviço a adquirir haverá dois limiares de distância definiveis: um (DT) que constitui o máximo tolerável, acima do qual se pode dizer não haver serviço, e outro (DD) que corresponde ao máximo desejável em termos do que se poderia chamar um bom nível de acesso (e portanto de serviço à população).

Assim sendo, poderemos definir três níveis de serviço ao longo do eixo das distância (entre o utilizador e o serviço) conforme se assinala na fig.1.

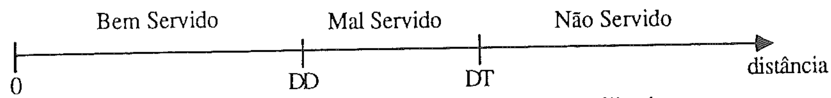


Fig.1 - Categorização do nível de serviço prestado aos utilizadores em função da distância de acesso.

Não sendo possível esperar que todos os cidadãos estejam em situação de rigorosa igualdade em termos de acesso a determinado tipo de equipamento (o que será viável com uma instalação desse equipamento na própria habitação de cada um dos cidadãos), será no entanto correcto procurar que no território em análise não haja grandes assimetrias no que respeita à categoria (das três referidas) em que as diferentes unidades de análise (p.ex. freguesias ou quadrículas) são classificáveis.

3 - Modelos Decisionais

A escassez de meios para investimento e este desejo de evitar grandes assimetrias recomendam por isso uma estratégia raramente seguida no nosso país mas que é fundamental para a resolução do problema: A consideração de uma escala móvel para os limiares de nível de serviço referidos (DD e DT).

A atitude tradicional tem sido no entanto a seguinte:

a) Analisa-se o problema como se não houvesse restrições de capital nem de recursos humanos e determina-se qual o limiar de distância (ou aliás de qualquer parâmetro) para que se possa dizer haver bom serviço.

b) Tentando de seguida implementar uma rede de equipamentos com esse nível de distância-limiar, verifica-se que tal não é possível por falta de recursos.

c) Determinam-se então "índices de carência" ou similares que mostram que os equipamentos são mais necessários onde há mais gente a residir e ordenam-se as potenciais localizações candidatas por ordem decrescente desse índice.

d) Face ao orçamento disponível, "corta-se" na lista assim produzida pela instalação cujo custo acumulado (desde a primeira) seja ainda comportável.

Como é conhecido de todos, este tipo de política leva sistematicamente à adopção pela Administração de um papel de "apaga-fogos" e à resignação das populações das zonas de menos densidade populacional e níveis de acesso aos equipamentos (comparados com as outras zonas) cada vez mais baixos. Não sendo a causa única do esvaziamento das zonas interiores do nosso país, não será concerteza dos menores obstáculos a qualquer política de regionalização e de repovoamento do interior que se queira (?) desenvolver.

A atitude que se recomenda é naturalmente muito diferente desta já que, ao considerar desde a primeira fase de análise o facto de os recursos serem finitos, o que vai ser móvel não vai ser a "faca" que corta na lista de prioridades (e deixa sistematicamente os mesmos de fora), mas sim o volume de investimentos disponíveis e como consequência o nível de serviço que é possível garantir a toda a população.

Usando um modelo matemático de optimização de expansão de capacidade de redes [1] pode responder-se em primeiro lugar à seguinte questão:

a) Face às limitações de recursos disponíveis, qual o valor mínimo praticável de DT, isto é, qual o valor máximo de distância que a partir de qualquer unidade residencial se terá que percorrer para ter acesso ao equipamento em causa ?

A resposta a esta questão será possível por resolução repetida do mesmo problema de optimização para vários valores de DT, obtendo-se para cada um desses valores o custo mínimo de expansão de rede actual até se chegar a uma configuração que garante essa distância máxima de acesso. A curva que se obtem tipicamente é da forma representada na fig.2.

O valor de DT poderá assim ser decidido tendo em atenção não só o nível de recursos disponíveis e as considerações funcionais sobre a distância de acesso do equipamento, mas também a eficácia marginal dos investimentos, devendo nomeadamente procurar-se optar por um nível de serviço (DT) como o do ponto B na fig.2, o qual foi atingido com pequenos custos e grandes melhorias de serviço a partir de C, e a partir do qual, em direcção a A, só com grandes custos se obtiveram ganhos significativos na acessibilidade ao serviço.

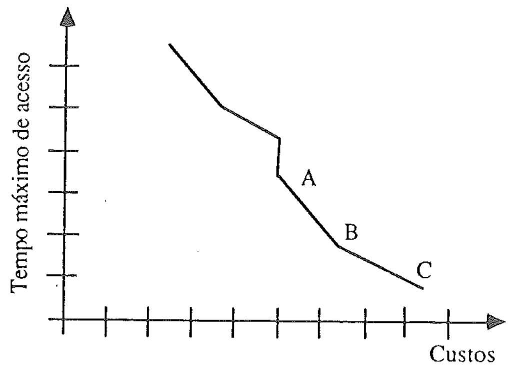
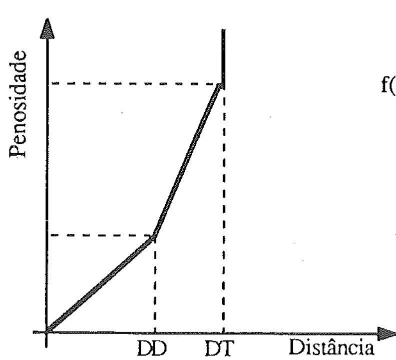


Fig.2 - Relação entre níveis de serviço garantido (distância máxima de acesso) e custos de investimento.

Escolhido o valor de DT, poder-se-á se for caso disso, passar à fase seguinte, em que se estipulam as áreas de influência de cada uma das instalações (actuais mais as resultantes do modelo da primeira fase) de modo a garantir a distribuição das populações tendo em atenção a sua capacidade e a minimização das distâncias a percorrer pelas populações.

Esta fase só tem sentido se se tratar de um serviço em que a Administração tem o direito de fixar a instalação a que o cidadão deve recorrer, havendo no entanto muitos outros em que o cidadão tem livre escolha da instalação. Nesses casos, a carga sobre cada instalação deve ser estimada através de modelos que tentem reproduzir os critérios do cidadão (por ex. gravitacionais ou de oportunidades).

Porque se admitiu a exisência de dois níveis de serviço em função da distância, é natural que se considerem de modo diverso os percursos feitos até DD e os acima de DD e abaixo de DT. Se imaginarmos uma função f(d) que represente a penosidade associada a um percurso d, teremos por exemplo, uma situação como a representsada na fig.3.



$$f(d)\begin{cases} d & \text{, se } d \leq DD \\ DD + W\,(d - DD) & \text{, se } DD < d \leq DT \end{cases}$$

Fig.3 - Função penosidade do percurso.

Assim será possível passar à resposta da seguinte questão:

b) Dados o máximo de investimento possível e a distância máxima a percorrer por qualquer utente DT (bem como a função penosidade do percurso), qual a configuração da rede de equipamentos que respeita aquelas restrições e garante o mínimo da penosidade média do acesso para o conjunto de todos os utentes ?

A resposta às duas questões aqui levantadas é já computacionalmente exequível mesmo em micro-computadores, pelo que não mais é argumento a sua dificuldade para que se evite a consideração das perguntas (incómodas ?).

Embora por limitações de espaço e a bem da clareza da exposição se tenha optado pela apresentação das questões em termos estatísticos, o mesmo tipo de análise é possível de execução em termos dinâmicos, caso em que se estuda todo o processo de transição da situação actual para uma outra de melhor serviço ao longo dos anos, considerando explicitamente as variações de procura e de disponibilidade de recursos.

Referências

[1] - José M. Viegas, 1984 • "Expansão de capacidade em Redes de Equipamentos: Metodologia e Aplicações", Dissertação para doutoramento em Engenharia Civil pelo Instituto Superior Técnico.

Investigação Operacional, Vol.7 nº1 (Junho/1987) 67 - 80

DUALIDADE E CONTABILIDADE ANALÍTICA

Jorge Guerreiro
Manuel Ramalhete

PETROGAL - E.P.

1 - Introdução

Esta comunicação insere-se na área da Contabilidade por centros de responsabilidade em empresas industriais. O objectivo consiste em apresentar uma abordagem nova ao problema da criação de incentivos aos gestores das diferentes secções principais e auxiliares; questão que entronca na problemática mais vasta da congruência de objectivos por parte dos vários responsáveis.

A abordagem que se apresenta baseia-se na Programação Linear e, por exploração dos resultados da Dualidade, procura dar resposta à determinação de custos e resultados previsionais por produtos e por secções.

Situando-se numa perspectiva de descentralização, a PL pode ainda orientar as decisões, nomeadamente de expansão, dos responsáveis das secções e fornecer critérios para os preços de transferência entre as várias secções e entre estas e a actividade comercial. Tal pode ser conseguido através duma adequada formalização do modelo de PL que determine o plano de produção, aparecendo a ligação à Contabilidade Analítica através da Dualidade.

Adopta-se a óptica de um processo de custeio variável, isto é, considera-se a imputação (sucessiva às secções e aos produtos) apenas dos custos variáveis. Esta posição pode justificar-se não só por ser preferível a imputação dos custos fixos de uma forma isolada e de acordo com critérios pré-definidos, como também pelo facto de numa situação de sub-utilização da capacidade produtiva existente a empresa poder ser levada a aceitar vender desde que o preço de venda dos produtos seja superior apenas aos custos variáveis.

Determinam-se simultaneamente os custos variáveis completos e os resultados por secções e por produtos associados ao plano de produção determinado pelo modelo. Saliente-se que os custos fixos por secções e por produtos podem ser obtidos através de matrizes de imputação que traduzam as inter-relações entre secções e entre estas e os produtos.

Será feita uma análise teórica por etapas, sendo posteriormente exemplificada através de uma caso prático.

2. Descrição do Modelo

Suponha-se uma empresa dedicada à produção e venda de vários produtos e organizada em secções auxiliares que prestam serviços às principais, concorrendo estas directamente para a produção dos produtos.

2.1 Caso clássico: o plano de produção é dado

Admita-se que a empresa possui s secções auxiliares, m secções principais e produz n produtos. A actividade das secções será medida em unidades de obra.

Seja $Y = (y_1,\ldots, y_s)$ o vector dos níveis de actividade das secções auxiliares. Estas secções fornecem serviços entre elas e às secções principais. Designe-se por c_{ij} a quantidade de unidades de obra da secção auxiliar i necessárias à produção de uma unidade de obra da secção auxiliar j.

Seja $Z = (z_1, \ldots, z_m)$ o vector dos níveis de actividade das secções principais e d_{ih} a quantidade de unidades de obra da secção auxiliar i necessárias à produção de uma unidade de obra da secção principal h.

Seja $X = (x_1, \ldots, x_n)$ o vector das quantidades a produzir dos diferentes produtos e b_{hk} a quantidade de unidades de obra da secção principal h necessárias à produção de uma unidade do produto k.

Tem-se então:

PRESTAÇÕES DAS SECÇÕES AUXILIARES

$$Y = CY + DZ \Leftrightarrow (I - C)Y - DZ = 0 \qquad (1)$$

em que

$$C = \left[c_{ij}\right] \quad \begin{matrix} i=1,\ldots,s \\ j=1,\ldots,s \end{matrix}$$

$$D = \left[d_{ih}\right] \quad \begin{matrix} i=1,\ldots,s \\ h=1,\ldots,m \end{matrix}$$

PRESTAÇÕES DAS SECÇÕES PRINCIPAIS

$$Z = BX \Leftrightarrow Z - BX = 0 \qquad (2)$$

em que

$$B = \left[b_{hk}\right] \quad \begin{matrix} h=1,\ldots,m \\ k=1,\ldots,n \end{matrix}$$

Conhecido X (plano de produção [(1)]) determina-se Z, e conhecido este determina-se Y:

$$Z = BX$$
$$(I - C)Y - DZ = (I - C)Y - DBX = 0$$
$$Y = (I - C)^{-1} DBX \qquad (3)$$

A partir daqui determinam-se facilmente os custos variáveis completos das unidades de obra das secções, os custos dos produtos e os resultados respectivos. Com efeito, seja:

$V = (v_1, \ldots, v_n)$ o vector dos preços de venda dos produtos;
$G = (g_1, \ldots, g_n)$ o vector dos custos variáveis directos dos produtos (matérias primas, mão-de-obra directa,...);
$A = (a_1, \ldots, a_s)$ o vector dos custos variáveis directos das secções auxiliares;
$P = (p_1, \ldots, p_m)$ o vector dos custos variáveis directos das secções principais.

Designando por

$\alpha = (\alpha_1, \ldots, \alpha_s)$ o vector dos custos variáveis completos das unidades de obra das secções auxiliares,
$\beta = (\beta_1, \ldots, \beta_m)$ o vector dos custos variáveis completos das unidades de obra das secções principais,
$\gamma = (\gamma_1, \ldots, \gamma_n)$ o vector dos custos variáveis completos dos produtos,
$L = (l_1, \ldots, l_n)$ o vector das margens brutas unitárias dos produtos,

tem-se:

$$\alpha = A + C'\alpha \Leftrightarrow \alpha = (I - C')^{-1}A \qquad \text{(Secções Auxiliares)} \quad (4)$$
$$\beta = P + D'\alpha \Leftrightarrow \beta = P + D'(I - C')^{-1}A \qquad \text{(Secções Principais)} \quad (5)$$
$$\gamma = G + B'\beta \Leftrightarrow \gamma = G + B'P + B'D'(I - C')^{-1}A \qquad \text{(Produto)} \quad (6)$$
$$L = V - \gamma \Leftrightarrow L = V - G - B'P - B'D'(I - C')^{-1}A \qquad \text{(Produto)} \quad (7)$$

A obtenção dos custos variáveis globais das secções e dos produtos, bem como a margem bruta total não reveste dificuldade.

A conclusão a retirar é que nesta óptica - caso clássico - *a Margem Bruta Total é exclusivamente ventilada pelos produtos finais, pois os preços de transferência entre secções e destas aos produtos coincidem com os custos variáveis completos respectivos.*

---

(1) - Não se inclui a variação de stocks para facilidade de exposição.

2.2 Modelo de empresa englobando secções principais e auxiliares

2.2.1 Restrições de mercado

O modelo de Programação Linear que se apresenta traduz o objectivo de maximização Assim, o modelo de empresa reveste a seguinte forma:

Maximizar MBT = $(V - G)X - AY - PZ$ (8)

sujeito a

(prestações das secções auxiliares) $(I - C)Y - DZ = 0$ (9)

(prestações das secções principais) $Z - BX = 0$ (10)

(mercado) $X \leq M$; $X, Y, Z \geq 0$ (11)

a que corresponde o problema dual

Minimizar $w = M\pi$ (12)

sujeito a

$$- B'\mu + \pi \geq V - G \quad (13)$$

$$(I - C)'\lambda \geq - A \quad (14)$$

$$- D'\lambda + \mu \geq - P \quad (15)$$

$$\lambda, \mu \text{ livres; } \pi \geq 0$$

em que:

$\lambda = (\lambda_1, \ldots, \lambda_s)$ é o vector das variáveis duais associadas às secções auxiliares,

$\mu = (\mu_1, \ldots, \mu_m)$ é o vector das variáveis duais associadas às secções principais,

$\pi = (\pi_1, \ldots, \pi_n)$ é o vector das variáveis duais associadas aos produtos.

Atendendo à Propriedade dos Desvios Complementares:

SECÇÕES AUXILIARES

Se o plano de produção obrigar, como será natural, ao funcionamento das secções

$$(I - C)'\lambda = -A$$

ou, o que é equivalente,

$$\lambda = - (I - C')^{-1}A \quad (16)$$

Conclusão - A menos de sinal, as variáveis duais associadas às secções auxiliares coincidem com os custos variáveis completos das unidades de obra respectivas determinados no caso clássico (4), isto é, $\lambda = -\alpha$.

SECÇÕES PRINCIPAIS

De forma análoga tem-se

$$Z > 0 \Leftrightarrow - D'\lambda + \mu = - P \Leftrightarrow$$

$$\Leftrightarrow \mu = - P + D'\lambda$$

e atendendo a (16), vem

$$\mu = - \left[P + D' (I - C')^{-1}A\right] \quad (17)$$

Donde se que $\mu = -\beta$, isto é, as variáveis duais associadas às secções principais coincidem a menos de sinal com os custos variáveis completos das unidades de obra respectivas determinadas no caso clássico (5).

PRODUTOS

Atendendo à 1ª restrição do dual (13), vem

$$\pi \geq V - (G - \beta'\mu)$$

$$\pi \geq V - \left[G + B'P + B'D'(I - C')^{-1}A\right],$$

o que permite concluir que:

Por um lado, se o plano óptimo contemplar a produção do produto i ($x_i > 0$), então $\pi_i = v_i - \lambda_i$ ou, o que é equivalente, a variável dual associada à restrição de mercado de um produto a ser produzido coincide com a margem bruta unitária desse produto.

Por outro lado, se para algum produto se verificar $\pi_i > v_i - \gamma_i$, então $x_i = 0$ (a sua produção não deve ser contemplada), mas nesse caso, pelo facto de $x_i = 0 < m_i$, tem-se ainda, pela Propriedade dos Desvios Complementares, que $\pi_i = 0$, isto é, $v_i < \gamma_i$, o que significa que esse produto tem margem bruta unitária negativa ($v_i - \gamma_i < 0$) não sendo, logicamente, contemplada a sua produção. Sendo δ_j a variável desvio dual associada (custo reduzido), verifica-se que $\delta_i = \gamma_i - v_i$, sendo portanto o simétrico da margem bruta unitária (negativa).

Verifica-se, portanto, que a Margem Bruta Total, tal como no caso clássico, é exclusivamente ventilada pelos produtos. Pode mesmo dizer-se que há uma imputação da Margem Bruta Total exclusivamente à actividade comercial, pois são as restrições de mercado activas que determinam as margens brutas unitárias positivas ($\pi_i > 0$).

Em síntese:

1. Pode afirmar-se que a abordagem clássica coincide com a abordagem da
2. A resolução do modelo de empresa apresentado permite obter em simultâneo o
3. Importa ainda salientar que numa perspectiva de descentralização os custos
4. Este modelo apresenta ainda a vantagem, de facilitar a análise da possibilidade de alargamento das restrições de mercado e, em termos económicos, até onde interessa avançar, por exemplo, numa estratégia de aumento das vendas.

2.2.2 Restrições de capacidade das secções auxiliares e principais

A inclusão no modelo de restrições de capacidade nas secções auxiliares e principais não se reveste de dificuldade. O problema primal assume então a seguinte forma:

Maximizar MBT = $(V - G)X - AY - PZ$ (18)

sujeito a

(prestações das secções auxiliares) $(I - C)Y - DZ = \dot{0}$ (19)

(prestações das secções principais) $Z - BX = 0$ (20)

(capacidade das secções auxiliares) $Y \leq S$ (21)

(capacidade das secções principais) $Z \leq R$ (22)

(mercado) $X \leq M$; $X, Y, Z \geq 0$ (23)

a que corresponde o problema dual

Minimizar $w = S\eta + R\rho + M\pi$ (24)

sujeito a

$$- B'\mu + \pi \geq V - G \quad (25)$$

$$(I - C)'\lambda + h \geq - A \quad (26)$$

$$- D'\lambda + \mu + \rho \geq - P \quad (27)$$

$$\lambda, \mu \text{ livres; } \eta, \rho, \pi \geq 0 \quad (28)$$

em que:

$\eta = (\eta_1, \ldots, \eta_s)$ é o vector das variáveis duais associadas às restrições de capacidade das secções auxiliares,

$\rho = (\rho_1, \ldots, \rho_m)$ é o vector das variáveis duais associadas às restriçoes de capacidade das secções principais,

λ, μ e π têm o significado já referido em 2.2.1.

Tal como anteriormente, atendendo à Propriedade dos Desvios Complementares:

SECÇÕES AUXILIARES

$$Y > 0 \Leftrightarrow (I - C')\lambda + \eta = - A \Leftrightarrow$$

$$\Leftrightarrow \lambda = - \left[(I - C')^{-1}A + (I - C')^{-1}_{\eta}\right] \quad (29)$$

$$l = - \left[(I - C')^{-1}A + C^{*}\eta + \eta\right],$$

em que $C^{*} = (I - C')^{-1} - I$

Isto é, a menos de sinal, tem-se:

preço das unidades de obra das secções auxiliares	=	custo variável completo das unidades de obra	+	margens das secções auxiliares respectivas

em que o custo variável completo inclui a margem das secções auxiliares fornecedoras.

Verifica-se assim que, ao contrário do caso anterior (16), este modelo permite ventilar a Margem Bruta Total pelas Secções Auxiliares, uma vez que η é afinal o vector das margens unitárias das Secções Auxiliares.

De salientar que as Secções Auxiliares com capacidade disponível têm margem nula e, nesse caso, o preço de tranferência coincide com o custo variável completo.

SECÇÕES PRINCIPAIS

$$Z > 0 \Leftrightarrow - D'\lambda + \mu + \rho = -P \Leftrightarrow$$

$$\Leftrightarrow \mu = -P + D'\lambda - \rho,$$

e atendeno a (29), vem:

$$\mu = - \left[P + D'(I - C')^{-1}A + D'(I - C')^{-1}\eta + \rho\right] \qquad (30)$$

Donde se conclui que, a menos de sinal, se tem:

preço das unidades de obra das secções principais	=	custo variável completo das unidades de obra	+	margens das secções principais respectivas

em que o custo variável completo inclui também a margem das secções auxiliares, já que a transferência se processa a um preço que inclui a margem da secção fornecedora.

Ao contrário do resultado expresso em (17), o preço das unidades de obra das Secções Principais inclui a margem unitária respectiva (ρ), para além do custo variável completo contemplar a margem das secções fornecedoras $\left[D'(I - C')^{-1}\eta\right]$. Este modelo permite então a imputação às Secções Principais da Margem Bruta Total.

Também neste caso se verifica que as Secções Principais com capacidade disponível têm margem nula e, consequentemente, o preço de transferência coincide com o custo variável completo.

Consequentemente, esta abordagem permite, do ponto de vista duma gestão descentralizada, criar os incentivos correctos - no sentido de congruência com o objectivo geral da empresa - a uma expansão da capacidade produtiva no caso de esta se encontrar saturada.

PRODUTOS

Tal como no modelo anterior, se o plano óptimo contemplar a produção do produto i (xi > 0), então a avriável dual associada à restrição de mercado respectiva coincide com a margem bruta unitária desse produto, isto é,

$$x_i > 0 \Leftrightarrow \pi_i = v_i - \gamma_i$$

tem-se pois:

Preço de venda do produto	=	Custo variável completo do produto	+	Margem bruta unitária do produto

Por outro lado, para os produtos em que se verificar $\pi_i > v_i - \gamma_i$, a sua produção não é contemplada ($x_i = 0$), pois que, como se viu atrás, este têm Margem Bruta Unitária negativa, dada neste caso pelos custos reduzidos, δ_i, a menos de sinal. Isto é,

$$\delta_i = \gamma_i - v_i \Leftrightarrow x_i = 0 \qquad (v_i < \gamma_i)$$

Em síntese:

1) Esta abordagem engloba a abordagem anterior como caso particular, coincidindo os resultados obtidos em ambas desde que nenhuma das restrições de capacidade (das Secções Auxiliares e Principais) se encontre esgotada;
2) A solução dual deste modelo de empresa indica os preços de transferência das unidades de obra das Secções Auxiliares e Principais, os quais incluem para além dos custos variáveis completos as margens unitárias respectivas. Para além disso, tal como no modelo anterior, fornece informação relativa ao custo variável completo dos produtos e respectiva margem unitária.

3) Neste caso a Margem Bruta Total (MBT) é imputada às Secções Auxiliares,

$$MBT = S\eta + R\rho + M\pi,$$

enquanto que no caso anterior a imputação recaía apenas nos produtos, isto é,

$$MBT = M\pi$$

4) Este modelo tem a vantagem de alargar a perspectiva anterior de análise e contribuir assim para uma política adequada de incentivos nas áreas Comercial (Produtos) e Produtiva (Secções Auxiliares e Principais) por via do estabelecimento de preços dos produtos e preços de transferência das unidades de obra congruentes com a política óptima de expansão da empresa.

3. Aplicação a um Caso (1)

3.1. Enunciado

Uma empresa dedica-se à produção de óleos de origem vegetal. As matérias-primas principais utilizadas são o Coconote, o Coco, a Soja e o Girassol. Para além destas, a empresa utiliza ainda duas matérias-primas secundárias: Soda Cáustica e Terra Activada. As matérias-primas principais são sujeitas às operações de Preparação (incluindo Moagem), Prensagem e Refinação, obtendo-se finalmente, para cada tipo de matéria-prima, um produto principal - o óleo refinado - e dois sub-produtos - farinha (para rações) e ácidos gordos.

As secções fabris da fábrica de óleos são as seguintes:

- Armazém de Matérias-Primas, onde são armazenadas as oleaginosas atrás indicadas;
- Preparação, onde as oleaginosas são trituradas e farinadas;
- Prensagem, onde os produtos obtidos na secção anterior são espremidos em prensas e se faz a separação do óleo da respectiva farinha;
- Refinação, secção onde os óleos obtidos na secção anterior são tratados com
- Armazém de óleos, onde os óleos refinados são armazendos e preparados para expedição.

Além destas secções, ligadas directamente à produção, encontram-se ainda na fábrica, com custos variáveis, as seguintes secções de apoio:

- Central de Vapor, onde é produzido o vapor necessário à fábrica de óleos e a
- Oficina, que realiza a maior parte dos trabalhos de conservação e que colabora
- Trabalhos Diversos, grupo de homens que se dedica aos mais variados trabalhos tais como cargas e descargas, limpeza da fábrica, etc.

Existem ainda, exclusivamente com custos fixos, as seguintes secções:

- Armazéns Gerais, onde são armazenados os prosutos não incluídos nos armazéns já indicados;
- Serviços Gerais, que abrange todos os restantes serviços da fábrica, como a Direcção Técnica, escritório fabril, guardas, e pessoal do refeitório e balneários, etc.

Nesta empresa a *Contabilidade Industrial* está organizada de acordo com o critério das *secções homogéneas*, sendo as respeitantes a secções com custos variáveis e as respectivas unidades de obra as seguintes:

- Armazém de Matérias-Primas: tonelada movimentada (ton.);
- Preparação: hora-máquina (H-M);
- Prensagem: hora-máquina (H-M);
- Refinação: hora-máquina (H-M);
- Armazém de óleos: tonelada movimentada (ton.);
- Central de Vapor: tonelada de vapor produzida (ton.);
- Oficina: hora-homem (H-H);
- Trabalhos Diversos: hora-homem (H-H).

Os dados técnico-económicos relevantes para o planeamento da sua actividade no

(1) Retirado da obra *"Programação Linear"*, vol.I, de Ramalhete, Guerreiro e Magalhães, McGraw-Hill, 1984. Lisboa.

Fornecimentos das Secções Auxiliares por unidade de obra

Fornecedores \ Clientes	Armazém Matérias-Primas	Preparação	Prensagem	Refinação	Armazém Óleos	Central Vapor	Oficina
Central Vapor			.42	.8	.27		.005
Oficina	.015	.4	.4	.5	.1	.1	
Trab.Div.	.5		.9		1.15		

Prestações fixas das Secções Auxiliares

Central de Vapor	40
Oficina	60
Trabalhos Diversos	150

Prestações das Secções Principais/Tonelada Produzida

Secções \ Produtos	Óleo Palmiste	Óleo Coco	Óleo Soja	Óleo Girassol
Armazém M-P	3	1.75	2.5	2
Preparação	1	1.75	.75	1.25
Prensagem	2	1.5	2.5	1.5
Refinação	1.5	1.25	1.25	1.75
Armazém Óleos	1	1	1	1

Capacidades Mensais das Secções

Central Vapor	1 000 ton.	Arm.Mat.-Primas	3 500 ton.
Oficina	1 400 H-H	Preparação	700 H-M
Trab.Diversos	3 500 H-H	Prensagem	700 H-M
		Refinação	700 H-M
		Arm.Óleos	680 ton.

Consumo de Matérias-Primas Secundárias
e Produção Sub-Produtos/Tonelada Produzida

Produtos \ M-P Secundárias Sub-Produtos	Soda Cáustica	Terra Activada	Farinha	Ácidos Gordos
Óleo Palmiste	.042	.011	1.853	.20
Óleo Coco	.044	.012	.736	.12
Óleo Soja	.040	.010	1.400	.15
Óleo Girassol	.045	.013	.958	.10

Vendas Máximas Previstas

Óleo Palmiste	150 ton.
Óleo Coco	60 ton.
Óleo Soja	250 ton.
Óleo Girassol	300 ton.

Preços de Venda (10^3 Esc./Ton.)

Óleo Palmiste	50	Farinha Palmiste	12
Óleo Coco	40	Farinha Coco	8
Óleo Soja	45	Farinha Soja	10
Óleo Girassol	45	Farinha Girassol	10
		Ácidos Gordos	24

Custos MAtérias-Primas
(10^3 Esc./Ton.)

Coconote	15
Coco	13
Soja	14
Girassol	15
Soda Cáustica	10
Terra Activada	18

Custos Variáveis Directos das Secções (10^3 Esc./Unidade Obra)

Central Vapor	.210	Arm.Mat.-Primas	.015
Oficina	.010	Preparação	.090
Trab.Diversos	.045	Prensagem	.180
		Refinação	.075
		Armazém Óleos	.021

3.2. Formalização

VARIÁVEIS PRINCIPAIS

x_1 - produção (ton.) de Óleo de Palmiste
x_2 - produção (ton.) de Óleo de Coco
x_3 - produção (ton.) de Óleo de Soja
x_4 - produção (ton.) de Óleo de Girassol
y_1 - nível de actividade (ton.) da Central de Vapor
y_2 - nível de actividade (H-H) da Oficina
y_3 - nível de actividade (H-H) dos Trabalhos Diversos
z_1 - nível de actividade (ton.) do Armazém de Matérias-Primas
z_2 - nível de actividade (H-M) da Preparação
z_3 - nível de actividade (H-M) da Prensagem
z_4 - nível de actividade (H-M) da Refinação
z_5 - nível de actividade (ton.) do Armazém de Óleos

RESTRIÇÕES

Prestações das Secções Auxiliares

C.Vapor	$y_1 =$	$.005\,y_2 + .42\,z_3 + .8\,z_4 + .27\,z_5 + 40$
Oficina	$y_2 = .1y_1$	$+ .015\,z_1 + .4\,z_2 + .4\,z_3 + .5\,z_4 + .1\,z_5 + 60$
Trab.Div.	$y_3 =$	$.5\,z_1 + .9\,z_3 + 1.15\,z_5 + 150$

Prestações das Secções Principais

A.M. Primas	$z_1 =$	$3x_1 + 1.75\,x_2 + 2.5\,x_3 + 2\,x_4$
Preparação	$z_2 =$	$x_1 + 1.75\,x_2 + .75\,x_3 + 1.25\,x_4$
Prensagem	$z_4 =$	$2\,x_1 + 1.5\,x_2 + 2.5\,x_3 + 1.5\,x_4$
Refinação	$z_4 =$	$1.5\,x_1 + 1.25\,x_2 + 1.25\,x_3 + 1.75\,x_4$
Arm. Óleos	$z_5 =$	$x_1 + x_2 + x_3 + x_4$

Capacidades das Secções Auxiliares

Central de Vapor	$y_1 \leq 1\,000$
Oficina	$y_2 \leq 1\,400$
Trab.Diversos	$y_3 \leq 3\,500$

Capacidades das Secções Principais

Arm.Mat.-Primas	$z_1 \leq 3\,500$
Preparação	$z_2 \leq 700$
Prensagem	$z_3 \leq 700$
Refinação	$z_4 \leq 700$
Arm. Óleos	$z_5 \leq 680$

Mercados

Óleo Palmiste	$x_1 \leq 150$
Óleo Coco	$x_2 \leq 60$
Óleo Soja	$x_3 \leq 250$
Óleo Girassol	$x_4 \leq 300$

Margens Brutas Directas dos Produtos

	Valor	Óleo Palmiste x_1		Óleo Coco x_2		Óleo Soja x_3		Óleo Girassol x_4	
	unitário	coef. unit.	valor	coef. unit.	valor	coef. unit.	valor	coef. unit.	valor
Produtos									
Óleo palmiste	50	1	50						
Óleo coco	40			1	40				
Óleo soja	45					1	45		
Óleo girassol	45							1	45
Sub-produtos									
Farinha palmiste	12	1.853	22.236						
Farinha coco	8			.736	5.888				
Farinha soja	10					1.400	14.00		
Farinha girassol	10							.958	9.58
Ácidos gordos	24	.20	4.8	.12	2.88	.15	3.6	.10	2.4
Mat.-Pr. princ.									
Coconote	15	3	45						
Coco	13			1.75	22.75				
Soja	14					2.5	35		
Girassol	15							2	30
Mat.-Prima sec.									
Soda cáustica	10	.042	.42	.044	.44	.040	.4	.045	.45
Terra activada	18	.011	.198	.012	.216	.010	.18	.013	.234
RESULTADO		31.418		25.362		27.02		26.296	

FUNÇÃO OBJECTIVO

Margem Bruta Total (MBT) = Receitas dos produtos e sub-produtos - Custos variáveis directos dos produtos - Custos variáveis directos das Secções Auxiliares - Custos variáveis directos das Secções Principais =

= Margem bruta directa dos produtos - Custos variáveis directos Secções Auxiliares - Custos variáveis directos Secções Principais

Atendendo ao quadro 2.10 - margens brutas directas dos produtos - e à informação sobre os custos variáveis directos das diferentes secções, tem-se para a FO:

$$\begin{aligned} \text{Maximizar MBT} = {} & 31.418\, x_1 + 25.362\, x_2 + 27.020\, x_3 + 26.296\, x_4 - \\ & - .21\, y_1 - .01\, y_2 - .045\, y_3 - \\ & - .015\, z_1 - .09\, z_2 - .10\, z_3 - .075\, z_4 - .021\, z_5 \end{aligned}$$

Em síntese, tem-se o modelo de PL:

Variáveis	x_1	x_2	x_3	x_4	y_1	y_2	y_3	z_1	z_2	z_3	z_4	z_5	≥ 0
FO Max.	31.418	25.362	27.020	26.296	.2	-.01	-.045	-.015	-.09	-.018	-.075	-.021	
Sec.aux.					1	-.005				-.42	-.8	-.27	= 40
					-.1	1		-.015	-.4	-.4	-.5	-.1	= 60
							1	-.5		-.9		-1.15	= 150
Sec.princ.	-3	-1.75	-2.5	-2				1					= 0
	-1	-1.75	-.75	-1.25					1				= 0
	-2	-1.5	-2.5	-1.5						1			= 0
	-1.5	-1.25	-1.25	-1.75							1		= 0
	-1	-1	-1	-1								1	= 0
Capacid. secun. aux.					1								≤ 1.000
						1							≤ 1 400
							1						≤ 3 500
Capacid. secund. princ.								1					≤ 3 500
									1				≤ 700
										1			≤ 700
											1		≤ 700
												1	≤ 680
Mercado	1												≤ 150
		1											≤ 60
			1										≤ 250
				1									≤ 300

3.3 Solução e interpretação

3.3.1 Solução

Primal		Dual (10^3esc)	
X_1 (Óleo Palmiste)	= 125 ton		$\pi_1 = 0$
X_2 (Óleo Coco)	= 60 ton		$\pi_2 = 0.91225$
X_3 (Óleo Soja)	= 0 ton		$\pi_3 = 0$
X_4 (Óleo Girassol)	= 240 ton		$\rho_4 = 0$
Y_1 (Central Vapor)	= 1000 ton	$\lambda_1 = -4.60649$	$\eta_1 = 4.39318$
Y_2 (Oficina)	= 1050 H-H	$\lambda_2 = -0.03303$	$\eta_2 = 0$
Y_3 (Trab.Diversos)	= 1749 H-H	$\lambda_3 = -0.04500$	$\eta_3 = 0$
Z_1 (Arm. M-P)	= 960 ton	$\mu_1 = -0.03800$	$\rho_1 = 0$
Z_2 (Preparação)	= 530 H-M	$\mu_2 = -0.10321$	$\rho_2 = 0$
Z_3 (Prensagem)	= 700 H-H	$\mu_3 = -12.10797$	$\rho_3 = 9.93953$
Z_4 (Refinação)	= 682.5 H-M	$\mu_4 = -3.77671$	$\rho_4 = 0$
Z_5 (Arm. Óleos)	= 425 ton	$\mu_5 = -1.31980$	$\rho_5 = 0$

3.3.2 Interpretação

SECÇÕES AUXILIARES

$$\left[\text{preço da unidade de obra}\right] = \left[\text{custo variável directo} + \text{custo variável indirecto}\right] + \left[\text{margem unitária}\right]$$

Central Vapor

$$\underset{-\lambda_1}{4.60649} = (\underset{a_1}{0.21} + \underset{c_{21}}{.1} \times \underset{(-\lambda_2)}{0.03303}) + \underset{\eta_1}{4.39318}$$

Oficina

$$\underset{-\lambda_2}{0.03303} = (\underset{a_2}{0.01} + \underset{c_{12}}{0.005} \times \underset{(-\lambda_1)}{4.60649}) + \underset{\eta_2}{0}$$

Trabalhos Diversos

$$\underset{-\lambda_3}{0.04500} = (\underset{a_3}{0.045} + 0) + \underset{\eta_3}{0}$$

Observação 1 - A parcela custo variável indirecto corresponde às prestações recíprocas entre secções auxiliares.

Observação 2 - A Central de Vapor transfere a um preço que inclui a margem unitária respectiva, uma vez que a sua capacidade é plenamente utilizada.

Observação 3 - Os preços das Unidades de Obra da Oficina e trabalhos diversos coincidem com os custos variáveis completos, uma vez que não trabalham a plena capacidade.

SECÇÕES PRINCIPAIS

$$\left[\text{preço da unidade de obra}\right] = \left[\text{custo variável directo} + \text{custo variável indirecto}\right] + \left[\text{margem unitária}\right]$$

Arm. Matérias Primas

$$\underset{-\mu_1}{0.03800} = (\underset{p_1}{0.015} + \underset{d_{21}}{0.015} \times \underset{(-\lambda_2)}{0.03303} + \underset{d_{31}}{0.5} \times \underset{(-\lambda_3)}{0.04500}) + \underset{\rho_1}{0}$$

Preparação

$$0.10321 = (0.09 + 0.4 \times 0.03303) + 0$$

Prensagem

$$12.10797 = (0.18 + 0.42 \times 4.60649 + 0.4 \times 0.03303 + 0.9 \times 0.04500) + 9.93953$$

Refinação

$$3.77671 = (0.075 + .8 \times 4.60649 + 0.5 \times 0.3303) + 0$$

Arm. Óleos

$$1.131980 = (0.021 + 0.27 \times 4.60649 + 0.1 \times 0.03303 + 1.15 \times 0.045) + 0$$

Observação - Os preços das Unidades de Obra das secções principais que não apresentam a capacidade esgotada coincidem com os custos variáveis completos. Apenas a prensagem que (com este plano de produção) vê a sua capacidade esgotada é que tem estabelecido um preço de transferência que engloba a margem unitária respectiva.

PRODUTOS $(X > 0)$

$$\begin{bmatrix}\text{Preço de}\\ \text{Venda}\end{bmatrix} = \begin{bmatrix}\text{Custo variável}\\ \text{directo}\end{bmatrix. + \begin{bmatrix}\text{Custo variável}\\ \text{indirecto}\end{bmatrix} + \begin{bmatrix}\text{Margem}\\ \text{unitária}\end{bmatrix}$$

Óleo Palmiste ($x_1 = 125$ ton)

$$\underset{v_1}{50} = (\underset{g_1}{18.582} + \underset{b_{11} \times (-\mu_1)}{3 \times 0.03800} + \underset{b_{21} \times (-\mu_2)}{1 \times 0.10321} + \underset{b_{31} \times (-\mu_3)}{2 \times 12.10797} + \underset{b_{41} \times (-\mu_4)}{1.5 \times 3.77671} + \underset{b_{51} \times (-\mu_5)}{1 \times 1.3198}) + \underset{\pi_1}{0}$$

Óleo Coco ($x_2 = 60$ ton)

$$40 = (14.638 + 1.75 \times 0.03800 + 1.75 \times 0.10321 + 1.5 \times 12.10797 + 1.25 \times 3.77671 + 1 \times 1.3198) + 0.91225$$

Óleo Girassol ($x_4 = 240$ ton)

$$45 = (18.704 + 2 \times 0.03800 + 1.25 \times 0.10321 + 1.5 \times 12.10797 + 1.75 \times 3.77671 + 1 \times 1.3198) + 0$$

PRODUTO $(X = 0)$

$$\begin{bmatrix}\text{Preço Venda}\end{bmatrix} < \begin{bmatrix}\text{Custo variável directo} + \text{Custo variável indirecto}\end{bmatrix}$$

Óleo Soja ($x_3 = 0$ ton)

$$45 < (17.98 + 2.5 \times 0.03800 + .75 \times 0.10321 + 2.5 \times 12.10797 + 1.25 \times 3.77671 + 1 \times 1.3198) = 54.46302$$

Perda de oportunidade = 54.46302 - 45 = 9.46302, logo o produto não é produzido

A Margem Bruta Total é imputada aos Produtos, Secções Auxiliares e Secções Principais. Com efeito:

$$\begin{aligned} MBT &= (150 \times 0 + 60 \times 0.91225 + 250 \times 0 + 300 \times 0) + (1000 \times 4.39318 + \\ &+ 1400 \times 0 + 3500 \times 0) + (3500 \times 0 + 700 \times 0 + 700 \times 9.93953 + \\ &+ 700 \times 0 + 680 \times 0) - (40 \times 4.60649 - 60 \times 0.03303 - 150 \times 0.04500) \\ &= \underset{\text{Merc.}}{54.735} + \underset{\text{Sec.Aux.}}{4\,393.18} + \underset{\text{Sec.Princ.}}{6\,957.671} - \underset{\text{Prest.Fixas}}{192.9914} = 11\,212.5946 \end{aligned}$$

Bibliografia

[1] - HORNGREN, C.T. - Cost Accounting: A. Managerial Emphasis, 4 th.ed., Prentice/Hall International, Inc., 1977.

[2] - LACAZE, D. - Prix duaux et Comptabilité Analytique, Paris, 1980.

[3] - LACAZE, D. - Théorie des prix et décentralizatiom des décisions, Éditions du C.N.R.S., Paris, 1975.

[4] - MARTEAU, G.; J. SCHEID - Comptabilité Analytique et Controle de Gestion, P.U.F., 1974.

[5] - RAMALHETE, M.; J. GUERREIRO; A. MAGALHÃES - Programação Linear, vol. I, McGraw-Hill, Lisboa, 1984.

ANEXO
Solução de Computador

```
ROWS
 N  VALUE
 E  E1
 E  E2
 E  E3
 E  E4
 E  E5
 E  E6
 E  E7
 E  E8
 L  E9
 L  E10
 L  E11
 L  E12
 L  E13
 L  E14
 L  E15
 L  E16
 L  E17
 L  E18
 L  E19
 L  E20
COLUMNS
    OLPALM    VALUE          31.41800   E4      -    3.00000
    OLPALM    E5      -       1.00000   E6      -    2.00000
    OLPALM    E7      -       1.50000   E8      -    1.00000
    OLPALM    E17             1.00000
    OLCOCO    VALUE          25.36200   E4      -    1.75000
    OLCOCO    E5      -       1.75000   E6      -    1.50000
    OLCOCO    E7      -       1.25000   E8      -    1.00000
    OLCOCO    E18             1.00000
    OLSOJA    VALUE          27.02000   E4      -    2.50000
    OLSOJA    E5      -        .75000   E6      -    2.50000
    OLSOJA    E7      -       1.25000   E8      -    1.00000
    OLSOJA    E19             1.00000
    OLGIRA    VALUE          26.29600   E4      -    2.00000
    OLGIRA    E5      -       1.25000   E6      -    1.50000
    OLGIRA    E7      -       1.75000   E8      -    1.00000
    OLGIRA    E20             1.00000
    CVAPOR    VALUE   -        .21000   E1           1.00000
    CVAPOR    E2      -        .10000   E9           1.00000
    OFICIN    VALUE   -        .01000   E1      -     .00500
    OFICIN    E2              1.00000   E10          1.00000
    TRDIV     VALUE   -        .04500   E3           1.00000
    TRDIV     E11             1.00000
    ARMMPR    VALUE   -        .01500   E2      -     .01500
    ARMMPR    E3      -        .50000   E4           1.00000
    ARMMPR    E12             1.00000
    PREPAR    VALUE   -        .09000   E2      -     .40000
    PREPAR    E5              1.00000   E13          1.00000
    PRENSA    VALUE   -        .18000   E1      -     .42000
    PRENSA    E2      -        .40000   E3      -     .90000
    PRENSA    E6              1.00000   E14          1.00000
    REFINA    VALUE   -        .07500   E1      -     .80000
    REFINA    E2      -        .50000   E7           1.00000
    REFINA    E15             1.00000
    ARMOLE    VALUE   -        .02100   E1      -     .27000
    ARMOLE    E2      -        .10000   E3      -    1.15000
    ARMOLE    E8              1.00000   E16          1.00000
RHS
    VETOR1    E1             40.00000   E2          60.00000
    VETOR1    E3            150.00000   E9        1000.00000
    VETOR1    E10          1400.00000   E11       3500.00000
    VETOR1    E12          3500.00000   E13        700.00000
    VETOR1    E14           700.00000   E15        700.00000
    VETOR1    E16           680.00000   E17        150.00000
    VETOR1    E18            60.00000   E19        250.00000
    VETOR1    E20           300.00000
ENDATA
```

SECTION 1 - ROWS

NUMBER	...ROW..	AT	...ACTIVITY...	SLACK ACTIVITY	..LOWER LIMIT.	..UPPER LIMIT.	.DUAL ACTIVITY
1	VALUE	BS	11212.59880	11212.59880-	NONE	NONE	1.00000
2	E1	EQ	40.00000	.	40.00000	40.00000	4.60649
3	E2	EQ	60.00000	.	60.00000	60.00000	.03303
4	E3	EQ	150.00000	.	150.00000	150.00000	.04500
5	E4	EQ	.	.	.	.	.03800
6	E5	EQ	.	.	.	.	.10321
7	E6	EQ	.	.	.	.	12.10797
8	E7	EQ	.	.	.	.	3.77671
9	E8	EQ	.	.	.	.	1.31980
10	E9	UL	1000.00000	.	NONE	1000.00000	4.39318-
11	E10	BS	1050.14930	349.85070	NONE	1400.00000	.
12	E11	BS	1748.74979	1751.25021	NONE	3500.00000	.
13	E12	BS	960.00033	2539.99967	NONE	3500.00000	.
14	E13	BS	529.99934	170.00066	NONE	700.00000	.
15	E14	UL	700.00000	.	NONE	700.00000	9.93953-
16	E15	BS	682.49918	17.50082	NONE	700.00000	.
17	E16	BS	424.99967	255.00033	NONE	680.00000	.
18	E17	BS	125.00099	24.99901	NONE	150.00000	.
19	E18	UL	60.00000	.	NONE	60.00000	.91225-
20	E19	BS	.	250.00000	NONE	250.00000	.
21	E20	BS	239.99868	60.00132	NONE	300.00000	.

SECTION 2 - COLUMNS

NUMBER	.COLUMNS	AT	...ACTIVITY...	..INPUT COST..	..LOWER LIMIT.	..UPPER LIMIT.	.REDUCED COST.
22	OLPALM	BS	125.00099	31.41800	.	NONE	.
23	OLCOCO	BS	60.00000	25.36200	.	NONE	.
24	OLSOJA	LL	.	27.02000	.	NONE	9.46301-
25	OLGIRA	BS	239.99868	26.29600	.	NONE	.
26	CVAPOR	BS	1000.00000	.21000-	.	NONE	.
27	OFICIN	BS	1050.14930	.01000-	.	NONE	.
28	TRDIV	BS	1748.74979	.04500-	.	NONE	.
29	ARMMPR	BS	960.00033	.01500-	.	NONE	.
30	PREPAR	BS	529.99934	.09000-	.	NONE	.
31	PRENSA	BS	700.00000	.18000-	.	NONE	.
32	REFINA	BS	682.49918	.07500-	.	NONE	.
33	ARMOLE	BS	424.99967	.02100-	.	NONE	.

Investigação Operacional, Vol.7 nº1 (Junho/1987) 81 - 89

MODELOS DE PLANEAMENTO/GESTÃO PARA OPTIMIZAÇÃO DAS NECESSIDADES DE VIATURAS E TRIPULAÇÕES NUMA REDE DE TRANSPORTES

J. Romão Eusébio
Lélio Amado

RODOVIÁRIA NACIONAL
Av. Columbano Bordalo Pinheiro, 86
1096 LISBOA CODEX

Resumo: São apresentados os modelos desenvolvidos na Rodoviária Nacional, com vista à determinação das necessidades mínimas e racionalização da utilização de viaturas e pessoal de movimento, nomeadamente:

- Modelo de Transfega para optimização dos percursos das viaturas, que permite ainda caracterizar a tipologia necessária da frota de uma rede de transportes, a sua dimensão e localização;
- Modelo de Programação Linear para determinação do efectivo mínimo de pessoal necessário numa rede, num dia de trabalho específico, permitindo ainda a determinação dos seus horários de trabalho;
- Modelo de Programação Linear para determinação do efectivo mínimo anual necesário, as suas folgas e as suas férias.

Abstract: Mathematical Models developed at Rodoviária Nacional are presented for optimizing the needs and use of buses and crews, namely:

- Transhipment Model for optimizing buses (quantity, types, location and routes);
- Linear Programming Model to find the minimum number of crews and their time duties in a specific workday;
- Linear Programming Model to determine the minimum annual number of crews needed, and the corresponding structure of days-off and holidays.

Keywords: Modelling, Optimization, Scheduling.

1. Introdução

Apresesntam-se neste documento três modelos desenvolvidos na RN para apoiar a elaboração de uma escala de serviço numa rede de transportes em que a oferta está fixada.

Começamos com um modelo desenvolvido em 1977 que serve de base ao programa *"CIRCUITO"* para aplicação no estabelecimento dos percursos das viaturas e na determinação do seu quantitativo e tipologia.

Seguidamente apresenta-se o modelo e programa consequente *"HORÁRIOS"* que é aplicado para determinar optimamente os efectivos a trabalhar num dia de movimento e os respectivos horários de trabalho.

Conhecido o efectivo necessário diariamente, é preciso optimizar o efectivo total a afectar à rede, atendendo a que os trabalhadores terão folgas e férias. O modelo e programa que apresentamos no final é aplicado na determinação deste efectivo total e da sua estrutura óptima de férias e folgas.

Para além destas aplicações standard dos programas, apresentam-se também as aplicações marginais que é possível efectuar, nomeadamente simulações de horários de trabalho, e análises de sensibilidade a alterações da oferta.

2. Programa Circuito

Definida uma oferta de transporte a ser assegurada numa rede, o modelo permite responder fundamentalmente às questões: número mínimo de viaturas necessárias; sua tipologia; seus percursos.

Cada viagem que constitui a oferta é identificada por linha em que se insere, hora e terminus de início e de destino e tipo de viaturas aconselháveis. Para além destes dados é necessário identificar a rede de estradas da região em estudo com a indicação das direcções possíveis e do tempo de percurso em vazio entre cada dois entroncamentos ou terminus contíguos.

O modelo assenta em conceitos básicos:

- **Associação. Viagem Precedente e Seguinte**

  As viagens i e j estão associadas e escreve-se (i -> j) quando a viatura que efectua a viagemi, efectua imediatamente a seguir a viagem j.

  i denomina-se a viagem precedente

  j denomina-se a viagem seguinte

- **Circuito**

  Considerando as duas seguintes associações (i -> j) (j -> k) em que a viagem i é seguinte a uma e precedente de outra.

  Então (i -> j) (j -> k) é um circuito em que i é a viagem inicial e k a viagem final.

  Um circuito poderá ser composto por qualquer número de associações ≥ 1. Um circuito representa um percurso de uma viatura.

- **Solução**

  É um conjunto de associações possíveis em que:

  - cada viagem terá quando muito apenas uma precedente;
  - cada viagem terá quando muito apenas uma seguinte.

- **Função penalizadora de uma solução**

  Cada associação é penalizada mais ou menos fortemente conforme o seu nível de inconveniência no sistema.

  No programa desenvolvido e numa utilização normal é possível penalizar, (através de parâmetros adequados) as mudanças de linhas, mudanças de

  Cada circuito é penalizado de uma forma normalmente pesada uma vez que o principal objectivo é minimizar o número de viaturas no sistema.

O modelo é do tipo transfega assim formalizado:

N = Número de viagens

$$\text{Min} \sum_{i=1}^{N+1} \sum_{j=1}^{N+1} c_{ij}\, x_{ij}$$

sujeito às seguinte restrições

$$\sum_{j=1}^{N+1} x_{ij} = 1 \qquad (\, i = 1,\ldots, N\,)$$

$$\sum_{i=1}^{N+1} x_{ij} = 1 \qquad (j = 1,\ldots, N)$$

$$\sum_{j=1}^{N+1} x_{N+1,j} = N$$

$$\sum_{i=1}^{N+1} x_{i,N+1} = N$$

$$x_{ij} \geq o \text{ e inteiro} \qquad (i = 1,\ldots, N+1;\ j = 1,\ldots, N+1)$$

em que:

$$\underset{i,j \neq N+1}{x_{ij}} = \begin{cases} 0, & \text{se a associção (i -> j) não pertence à solução} \\ 1, & \text{caso contrário} \end{cases}$$

$$\underset{j=1,\ldots,N}{x_{N+1,j}} = \begin{cases} 0, & \text{se a viagem já tem precedente} \\ 1, & \text{caso contrário ou seja é a viagem inicial de um circuito} \end{cases}$$

$$\underset{i=1,\ldots,N}{x_{i,N+1}} = \begin{cases} 0, & \text{se a viagem i tem seguinte} \\ 1, & \text{caso contrário ou seja é a viagem final de um circuito} \end{cases}$$

$\sum_{j=1}^{N} x_{N+1,j}$ = Total de viagens iniciais, ou seja, número de viaturas associadas à solução.

$\sum_{i=1}^{N} x_{i,N+1}$ = Número de viagens finais, ou seja, número de viaturas associadas à solução.

$c_{ij} = (K1 \times ML_{ij}) + (K2 \times TPV_{ij}) + (K3 \times MV_{ij})$
se a associação i -> for possível e i, j $\neq$ N+1; i $\neq$ j

$c_{ij} = \infty$ se a associação (i-> j) não for possível e i, j $\neq$ N+1 ou i = j
$c_{i,N+1}$ = IMPOS (i = 1,..., N) {
{ Por cada circuito a função penaliza
{ a solução em 2 × IMPOS
$c_{N+1,j}$ = IMPOS (j = 1,..., N) {
$c_{N+1,N+1} = 0$

$$ML_{ij} = \begin{cases} 0, & \text{se a viagem i pertence à linha da viagem j} \\ 1, & \text{caso contrário} \end{cases}$$

TPV_{ij} = Tempo de percurso em vazio entre o terminus final da viagem i e o inicial da viagem j
MV_{ij} = é tanto maior quanto a diferença de tipologia de viatura exigida pela viagem i e a viagem j

K1, K2, K3 e IMPOS são parâmetros definidos pelo utilizador e que permitem posicionar relativamente a diversos critérios para avaliação de uma boa solução.

A matriz do problema apresenta-se seguidamente admitindo que as viagens estão ordenadas por hora de início.

N = NÚMERO DE VIAGENS
VIAGEM SEGUINTE

		1	2	3	4	••••	N	N+1	
VIAGEM PRECEDENTE	1	X	c_{12}	c_{13}	c_{14}	••••	c_{1N}	IMPOS	1
	2	X	X	c_{23}	c_{24}	••••	c_{2N}	IMPOS	1
	3	X	X	X	c_{34}	••••	c_{3N}	IMPOS	1
	4	X	X	X	X		c_{4N}	IMPOS	1
	•	•	•	•	•	•••• • •	•	•	
	•	•	•	•	•	•••• • •	•	•	
	N	X	X	X	X	••••••	X	IMPOS	1
	N+1	IMPOS	IMPOS	IMPOS	IMPOS	••••••	IMPOS	0	
		1	1	1	1	••••••	1		N

Como utilização marginal o modelo pode ser utilizado:

- na simulação de alteração de horários das viagens;
- fixado o total de viaturas a lançar na rede, na determinação da sua tipologia e respectivos percursos.

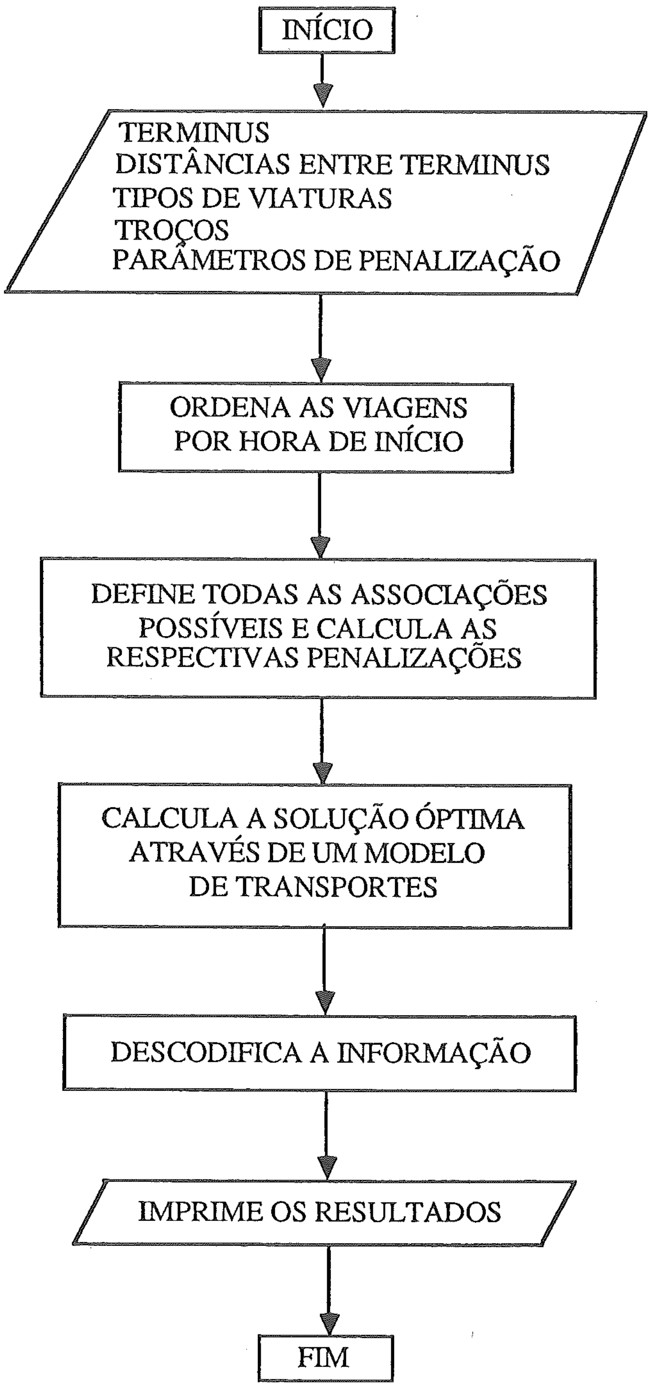
Este modelo tem limitações derivadas da sua manifesta "falta de memória" de que resultam consequências negativas quando estão em consideração viagens a exigirem diversificadas tipo de viaturas e/ou locais de pernoita.

Efectivamente quando se faz uma associação, não se leva em conta as viagens integrantes do circuito de que essa associação faz parte, mas tão somente as características da viagem precedente dessa associação.

Por estas razões os resultados do programa têm normalmente que ser ajustados normalmente sobretudo em zonas de movimento de características interurbanas.

Apresenta-se em esquema o fluxograma do programa circuito.

Programa Circuito



3. Programa Horários

Este programa aplica-se numa rede de transportes tipo urbana ou suburbana,,

- Tendo em consideração normas que dizem respeito ao horário de trabalho e tendo por objectivo a minimização dos custos de pessoal, qual o número mínimo de motoristas que garantem o serviço diário e os respectivos horários. Esta mesma questão é respondida para o grupo funcional cobradores.

As normas de horário de trabalho, bem como os elementos referentes a custos são

O programa começa por repartir o período de tempo com movimento na rede em fracções de 30 minutos ou de 1 hora consoante o interesse do utilizador. Em consequência, cada horário de trabalho apenas terá duas alternativas em relação a cada fracção de tempo: ou a preenche inteiramente, ou não a preenche totalmente. Obviamente que isto é uma simplificação da realidade e que se traduzirá em alguns problemas na implementação de soluções.

Numa segunda fase o programa calcula um conjunto de horários de trabalho

Aqueles horários terão que respeitar determinadas condições dadas pelas normas

O utilizador deverá assim fornecer elementos referentes a tempo de serviço normal,

O programa atribui a cada horário gerado um custo que é função de elementos fornecidos pelo utilizador e respeitante quer à utilização de um trabalhador adicional, quer aos critérios de atribuição de remunerações acessórias passando pelos custos unitários de remunerações por trabalho extraordinário e penalizações por trabalho nocturno e por refeição.

O programa permite ainda introduzir por processo especial horários de trabalho que não respeitem algumas das normas impostas.

Por cada uma das fracções em que o dia foi dividido, introduz-se o número de

Finalmente, e antes do processo de impressão final, o programa recorre ao modelo matemático de programação linear para, dentro dos horários criados, escolher um conjunto que minimize os custos de pessoal e que garanta as necessidades mínimas de pessoal apresentadas em cada fracção do dia.

O número de horários na solução dá o número de motoristas que garantem o serviço com custo global mínimo.

Apresenta-se seguidamente:

- Formulação matemática do problema;
- Fluxograma do programa;
- Exemplo de output.

Formulação Matemática

Pretende-se calcular,

$$x_i\,,\ i = 1, 2,\ldots, N$$

de forma a minimizar a F.O.

$$\text{Min} \sum_{i=1}^{N} c_i\, x_i$$

Respeitando as restrições

$$\sum_{i=1}^{N} a_{ij}\, x_i \geq b_j \qquad (\, j = 1,\ldots, M\,)$$

$$x_i = 0,\ 1,\ 2\ldots\ldots$$

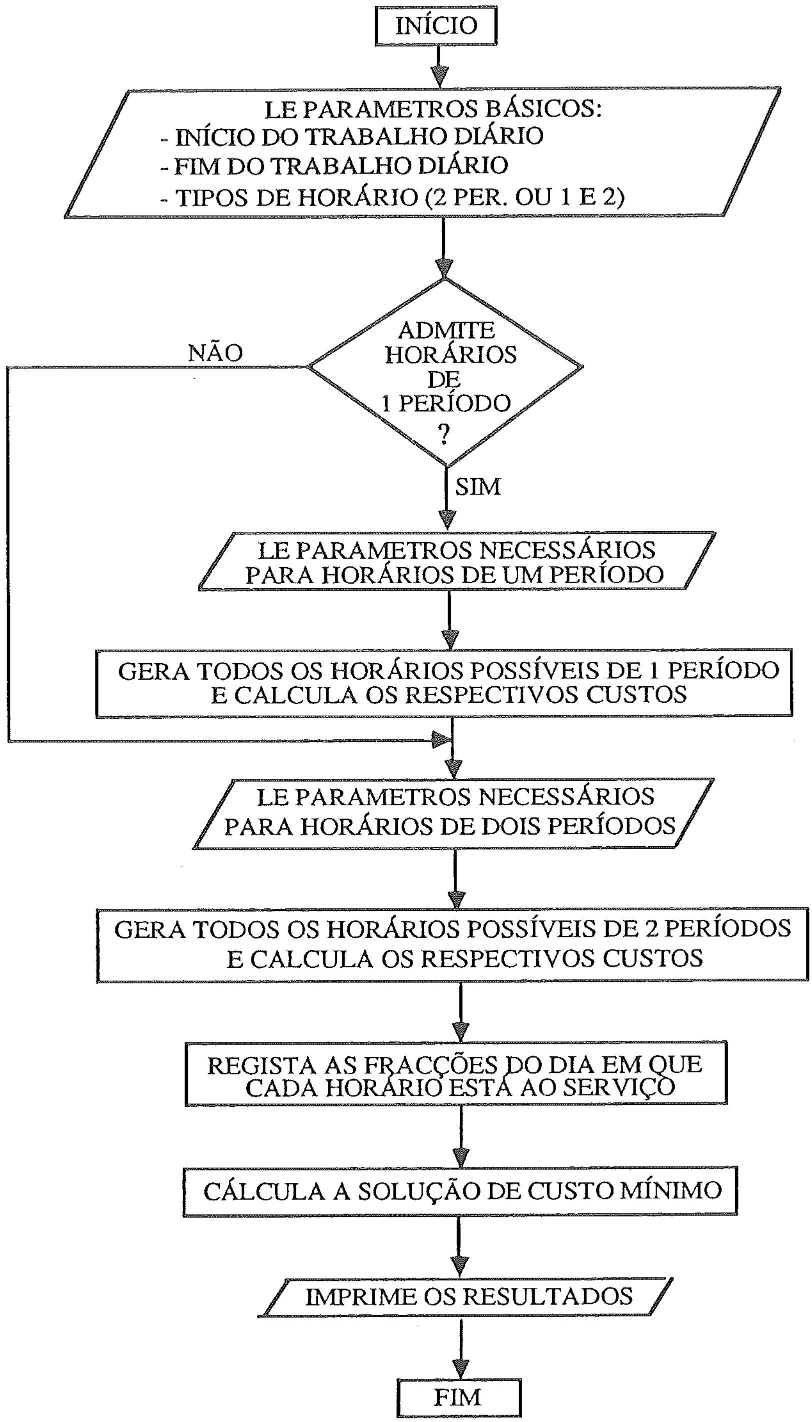
x_i é o número de motoristas com horário i
N é total de horários gerados
c_i = custo diário de um motorista com o hirário i
M = número de fracções em que o dia de trabalho foi repartido
b_j = número mínimo de motoristas necessário ao serviço durante a fracção de tempo j

$$a_{ij} = \begin{cases} 0, & \text{se o horário i não preenche a fracção de tempo j} \\ 1, & \text{caso contrário} \end{cases}$$

$$\sum_{i=1}^{N} a_{ij} = \text{Número de motoristas em serviço associado à solução } (x_i,\ i = 1,\ldots, N)$$

na fracção j.

Programa Horários



Exemplo de OUTPUT

A análise post optimal da solução é feita a partir dos valores duais associados a cada restrição e consiste fundamentalmente na pesquisa de melhores horários alternativos fora das normas introduzidas e na pesquisa das horas criticas onde valerá a pena um esforço especial na redifinição da oferta.

Este programa, introduzindo restrição adicional, pode também ser utilizado na determinação dos horários associados a uma solução em que, fixando o número de trabalhadores, são minimizadas as remunerações variáveis.

4. Programa SHEFF (Sistema de Horários, Escalas, Folgas e Férias)

Este programa surgiu da necessidade de, conhecendo as carências diárias de pessoal ao longo de semanas do ano, se dimensionar o pessoal de movimento e a respectiva estrutura de folgas e férias.

Parte dos pressupostos:

- Um trabalhador tem dois dias fixos seguidos de folga por semana;
- Um trabalhador tem um mês de férias seguido por ano (pressuposto ultrapassável na prática);
- Cada mês tem quatro semanas;
- Cada semana tem sete dias tipo, cada um eventualmente com necessidades diferentes de pessoal ao serviço;
- Em cada mês o número de trabalhadores necessários em cada dia tipo não varia com a semana;
- O ano tem doze meses.

Os dados a fornecer são apenas o número de trabalhadores necessários em cada dia tipo, em cada mês do ano.

Os resultados são:

- Total de trabalhadores necessários no ano;
- Total de trabalhadores com determinados dias de folga;
- Para cada grupo de trabalhadores com os mesmos dias de folga, a sua distribuição por meses de férias.

O problema é resolvido através da sua formulação como um problema de programação linear com (7 × 12) = 84 variáveis correspondentes às alternativas possíveis para um trabalhador quanto a dias de folga (7 alternativas) × Mês de férias (12 alternativas).

x_{ij} = Número de trabalhadores com dias de folga tipo i e mês de férias j
$(i=1,\ldots,7\ ;\ j=1,\ldots,12)$

$$\text{Min} \sum_{i=1}^{7} \sum_{j=1}^{12} x_{ij} \qquad \text{(Número de Trabalhadores)}$$

Restrições

$$\text{Min} \sum_{i=1}^{7} \sum_{j=1}^{12} a_{i,j,K,l}\, x_{ij} \geq b_{K,l} \quad ; \quad \begin{array}{l} K=1,\ldots,7 \text{ (7 dias da semana)} \\ l=1,\ldots,12 \text{ (12 meses)} \end{array}$$

$$x_{ij} \geq 0$$

$$a_{i,j,K,l} = \begin{cases} 0, & \text{se o trabalhador com dia de folga tipo i} \\ & \text{e mês de férias tipo j não está ao serviço} \\ & \text{no dia de semana K do mês l} \\ 1, & \text{caso contrário} \end{cases}$$

Como principal problema deste programa refira-se a necessidade de ajustar a solução continua a uma solução inteira. Este trabalho é realizado com apoio de programas em microcomputador e não ocupa mais do que meia hora a um técnico especializado.

5. Articulação do Programas Horários e Circuitos: Perspectivas Futuras

Como nota final a este programa refira-se que, definida a oferta, determinados o quantitativo, tipologia e percurso de viaturas, e calculado optimalmente os horários e efectivo de pessoal a movimentarem as viaturas, é necessário fazer a atribuição do serviço a cada trabalhador. Para apoiar este trabalho está em vias de inclusão um programa informático de manipulação de ficheiros.

Está no entanto nos nossos propósitos o desenvolvimento de modelos e programas por forma a realizar esta fase de forma mais automática, integrando os dois processos, isto é, o de optimização dos percursos das viaturas, e o da optimização dos serviços de pessoal.

Investigação Operacional, Vol.7 nº1 (Junho/1987) 91 - 107

DESCRIÇÃO DE UM MODELO DE EXPANSÃO DO SISTEMA ENERGÉTICO

Maria Natália Tavares
Victor Baptista

Direcção Central de Planeamento (EDP)

Resumo: Esta comunicação tem por objectivo a apresentação do modelo D.F.I., utilizado na análise e descrição do comportamento dos agentes económicos no sector energético, e que tem por base a teoria do equilíbrio económico geral.

A comunicação é constituída por 3 partes fundamentais:

- formulação matemática de um modelo de equilíbrio económico geral;
- interacções entre as diversas actividades de produção e consumo do sector energético representadas por um grafo orientado sem circuitos;
- algoritmo utilizado para a obtenção de um sistema de preços (e/ou quantidade) de um equilíbrio (solução numérica do problema).

Abstract: The goal of this paper is the description of an energy-economy model (DFI) where, based on general equilibrium theory, the problem behaviour of several economic agents of the sector are modelled.

This paper consists of three major components:

- mathematical formulation of a general economic equilibrum model;
- process network methodoly to represent the interrelations between producers and consumers;
- algorithm to obtain an equilibrium price system.

Keywords: Energy, Economics, Utility Functions, Multidecisional Problem, Nash Equilibrium.

1. Introdução

Tirando partido dos desenvolvimentos ocorridos nas diferentes áreas directamente ligadas ao processo de modelização, como por exemplo a programação matemática, a fiabilidade dos sistemas, a estatística, a análise das séries temporais e a teoria dos sistemas, tem-se assistido, a partir da década de 60, à implementação de modelos de complexidade crescente com vista a auxiliar os processos de tomada de decisão nos diversos sub-sectores da energia.

O objectivo de tais modelos é a expansão e gestão óptima dos recursos do subsector em estudo (por exemplo : expansão óptima do sistema electroprodutor), admitindo conhecidas todas as condições de fornteira (consumo, preços, índices técnico-económicos,...). Pelo seu isolamento relativo aos outros subsectores,tais modelos ignoram as inter-relações existentes entre eles, como a consideração de alternativas para a satisfação de determinada procura (por exemplo : o aquecimento ambiente no subsector residêncial serviços) tendo em conta as possibilidades de substituição entre as diversas formas de energia.

Estas necessidades de análise do sector energético como um todo e de implementação de novas políticas para fazerem face aos preços sempre crescentes do petróleo e à esgotabilidade das suas reservas, vieram originar os modelos energéticos, em grande expansão na década de 70, que têm por objectivo as interacções entre os diversos modelos de oferta e procura, com base nos modelos de crescimento económico (fig. 1).

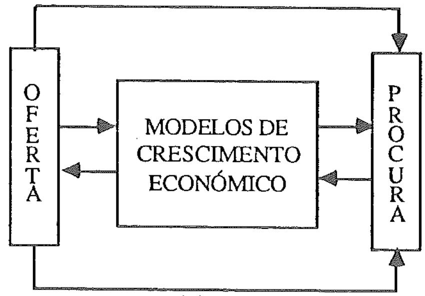


Fig. 1

Um modelo cujo objectivo é ajudar o processo de tomada de decisão no sistema energético global, terá de reflectir o caracter multidecisional do problema, identificando os seus elementos fundamentais. Esta identificação leva à definição dos agentes de decisão que o constituem (consumidores e produtore de energia) e à conceptualização adequada das inter-relações entre esses elementos, tendo em conta os seus objectivos, as suas incertezas e as suas preferências.

Um modelo deste tipo deve ter a flexibilidade suficiente para permitir, sempre que necessário, a restruturação do sistema, alteração do nível de detalhe de cada subsistema, das suas fronteiras, das suas condições iniciais, dos seus parâmetros, e ainda do tipo de comportamento de cada um dos agentes de decisão neles representados.

Em 1973, o Stanford Research Institute desenvolveu um modelo energético para analisar a estratégia a tomar pela Gulf Oil Corporation relativamente aos combustíveis sintéticos. Posteriormente, a Decision Focus Incorporated (DFI) integrou-o num modelo de energia global, tendo por base a teoria do equilíbrio económico geral, devidamente apoiado numa linguagem de computador flexível, dando origem ao modelo actual, DFI - Energy Economy Modeling System.

No âmbito dos estudos do PEN (Plano Energético Nacional) determinou-se a expansão optimizada do sistema energético a partir das procuras globais da energia correspondentes aos cenários de desenvolvimento da economia, utilizando um modelo dinâmico capaz de representar com detalhe suficiente os diferentes subsectores, o DFI Generalized Equilibrium Modeling Systems - GEMS.

2. Modelos de Comportamento Individual de Alguns Agentes Económicos

Este capítulo destina-se a expor modelos que pretendem representar as regras de comportamento individual de alguns agentes de decisão, típicos de economia. Não sendo possível no âmbito desta comunicação uma análise exaustiva do significado económico das propriedades desses modelos, limitar-nos-emos à sua formulação e à inidcação de algumas propriedades matemáticas relevantes, quer do ponto de vista económico, quer do ponto de vista de uma metodologia para identificação, estimação e validação desses modelos em face dos dados observáveis no comportamento real dos respectivos agentes de decisão.

Suponhamos uma economia ℓ bens de consumo, também designados por mercadorias. Seja $\Omega = \mathbb{R}_+^\ell$ o espaço das mercadorias $P = \mathbb{R}_+^\ell - \{0\}$ o espaço do sistema de preços. Designa-se por valor do vector de mercadorias $x \in \Omega$ ao sistema de preços $p \in P$ o número real não negativo px.

Admitamos ainda que existem M consumidores e N produtores dessas mercadorias, constituindo M+N agentes de decisão do sistema económico em estudo.

2.1 O consumidor maximizador de uma utilidade

A cada consumidor associamos um conjunto $\Delta \subset \Omega$ designado por um conjunto de planos de consumo admissíveis que supomos ser não vazio, fechado, convexo, inferiormente limitado e sobre o qual está definida uma relação binária $\geq$, dita de preferência, que é reflexiva, transitiva e completa, isto é, $\geq$ é uma pré-ordem completa de Δ.

Sejam x e y dois planos de consumo; diz-se que x é preferido a y se $x \geq y$, estritamente preferido se $x > y$ indiferente ou isopreferido se $x > y$ e $y > x$ (e escreve-se $x \iota y$). É claro que a relação de indiferença entre os planos de consumo é uma relação de equivalência, induzindo em Δ uma partição em classes de isopreferências.

A relação de preferência diz-se representável pela função utilidade

$$U: \Delta \rightarrow \mathbb{R}$$

se

$$U(x) \geq U(y) \Leftrightarrow x \geq y$$

e é claro que se $f : R \rightarrow \mathbb{R}$ é uma função crescente também $v = f o U$ é uma função utilidade que representa $\geq$. Se U for contínua, a relação de preferência diz-se continuamente representável por U e, é claro que uma condição necessária para isso aconteça é que os conjuntos

$$L^{\geq}(x, U) = \{y \in \Delta \mid y \geq x\} = U^{-1}\left(\left[U(x), \infty \right[\right)$$
$$L^{\leq}(x, U) = \{y \in \Delta \mid y \geq x\} = U^{-1}\left(\left] -\infty, U(x) \right] \right)$$

sejam fechados em Δ qualquer que seja $x \in \Delta$. Debreu mostrou que esta condição também é suficiente no caso em que Δ é um espaço topológico com uma base numerável de abertos, o que certamente acontece quando D tem a topologia induzida da topologia usual de $\mathbb{R}^{\ell}$.

Suponhamos agora que o consumidor está de posse dum vector de mercadorias w*. A teoria do consumidor maximizador de preferência assenta no seguinte princípio de racionalidade: dado um sistema de preços p, o consumidor escolhe no conjunto Δ um plano de consumos x*, maximal para a relação de preferências $\geq$, sujeito à condição do seu valor, px*, não exceder o orçamento disponível w.

Isto é, x* é um vector do conjunto orçamental

$$\beta(p, w) = \{x \in \Delta \mid px \leq w\}$$

e é tal que $\forall x \in \beta(p, w)\ x^* \geq x$. Se $\geq$ for representável por uma função utilidade U o problema de escolher em $\beta(p, w)$ um vector x* maximal para $\geq$ é equivalente a escolher uma solução do problema

$$\max_{x \in \beta(p,w)} U(x) \tag{1}$$

que sabemos existir no caso de U ser contínua e $\beta(p, w)$ um compacto de Δ.

A partir de agora supõe-se que o conjunto dos planos de consumo é $\Delta = \mathbb{R}_{+}^{\ell}$ e que a relação de preferências $\geq$ é continuamente representável por uma função utilidade U.

Se U for quase-concava, isto é, os conjuntos de nível

$$L^{\geq}(x, U) = \{y \in \Delta \mid U(y) \geq U(x)\} \tag{2}$$

são convexos, então o conjunto dos pontos solução do problema (1) é um compacto convexo de $\mathbb{R}^{\ell}$ ficando assim definida a correspondência de procura Marshalliana (fig.2)

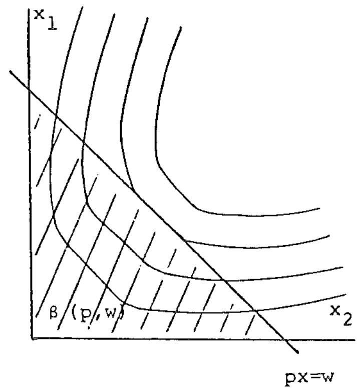


Fig. 2

$$X(p, w) = \{x^* \in b(p, w) \mid U(x^*) = \max_{x \in \beta(p,w)} U(x)\}$$

que é homogénea de grau 0 nos preços p e no orçamento w (isto é, $\forall \alpha > 0\ X(\alpha p, \alpha w) = X(p,w)$) e é superiormente contínua para todo o $(p, w) \in P \times \mathbb{R}$ tal que $w > \min_{x \in \Delta} px$.

Se além disso U for estritamente quase-concava, isto é, os conjuntos de nível (2) estritamente convexos, o problema (1) tem uma única solução e a correspondência de procura é de facto uma função, a função de procura Marshalliana, homogénea de grau 0 e contínua para todo o $(p, w) \in P \times \mathbb{R}$ tal que $w > \min_{x \in \Delta} px$ (fig.3).

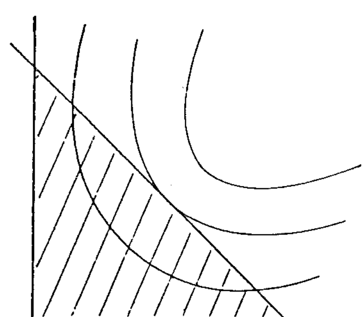


Fig. 3

Suponha-se a partir de agora que U é uma função quase-concava, estritamente crescente e de classe C^2.

Assim a função de procura Marshalliana X(p, w) é definida como a solução do problema

max	$u(x)$
sujeito a	$px \leq w$ $x \geq 0$

(3)

e a função V(p, w) = U(X(p, w)) designa-se por *utilidade indirecta* e representa a utilidade máxima que o consumidor pode atingir ao sistema de preços p e dispondo de uma orçamento w.

A monotonia estrita de U implica que a restrição orçamental é satisfeita como igualdade no ponto X(p, w) e portanto se $w > 0$, $p_i > 0$

$$\frac{\partial U}{\partial X_i} = \lambda(p, w)\, p_i \tag{4}$$

onde $\lambda(p, w) > 0$ é o multiplicador de Lagrange associado à restrição $px < w$.

A *função procura compensada* (Hicks) $\psi(p, u)$ é definida como a solução do problema

$$\boxed{\begin{array}{ll} \min & px \\ \text{sujeito a} & U(x) \geq u \\ & x \geq 0 \end{array}} \tag{5}$$

e a função $C(p, u) = p\, \psi(p, u)$ designa-se por *custo*.

X(p, w) e V(p, w) são homogéneas de grau 0 em (p, w), $\psi(p, u)$ é homogénea de grau o em p, C(p, u) é uma função positiva, crescente em U, crescente em p, positivamente homogénea em p, concava em p e, portanto, contínua em p.

Da quase concavidade estrita da função U conclui-se ainda que X e ψ são funções contínuas e que numa vizinhança dum sistema de preços $p \in \mathbb{R}^{\ell}_{+}$, onde $C(p, u) > 0$, C é diferenciável, estritamente crescente, estritamente concava na variável p e satisfaz a relação

$$\frac{\partial C}{\partial p_i}(p, u) = \psi_i(p, u)\ , \quad i = 1, \ldots, \ell \tag{6}$$

conhecido por *lema de Shephard.*

Pela monotonia estrita de U

$$\psi(p, w) = X(p, C(p, u)) \tag{7}$$

e fazendo $w = C(p, u)$ vem

$$\frac{\partial \psi_i}{\partial p_j}(p, u) = \frac{\partial X_i}{\partial p_j}(p, w) + \frac{\partial X_i}{\partial w}(p, w)\, \frac{\partial C}{\partial p_j}(p, u) \tag{7'}$$

e pelo lema de Shephard

$$\frac{\partial \psi_i}{\partial p_j}(p, u) = \frac{\partial X_i}{\partial p_j}(p, w) + \frac{\partial X_i}{\partial w}(p, w)\ X_j\ (p, u) \tag{8}$$

Se C for classe C^2, do lema de Shephard vem

$$K_{ij} = \frac{\partial \psi_i}{\partial p_j}(p, u) = \frac{\partial \psi_j}{\partial p_i}(p, u) = K_{ji} \tag{9}$$

isto é, a matriz de Slutsky $K = \nabla p\ y(p, u)$ é simétrica e a concavidade estrita da função custo C relativamente a p implica que K é semi-definida negativa.

Sendo $\psi(\circ, u)$ homogénea de grau 0 vem, pelo teorema de Euler,

$$\sum_{j=1}^{\ell} p_j\ \frac{\partial \psi_i}{\partial p_j}(p, u) = 0 \tag{10}$$

ou (por simetria de K)

$$\sum_{j=1}^{\ell} p_j\ \frac{\partial \psi_j}{\partial p_i}(p, u) = 0$$

o que é equivalente a:

$$\sum p_j \frac{\partial X_K}{\partial p_j}(p, w) + w \frac{\partial X_K}{\partial w}(p, w) = 0$$

De (8) e (9) conclui-se que K tem características $r(K) \leq \ell - 1$.

Da simetria de K e de (8) obtem-se a equação de Slutsky

$$\frac{\partial X_i}{\partial p_j}(p, w) + \frac{\partial X_i}{\partial w}(p, w) \,.\, X_j(p, w) + \frac{\partial X_j}{\partial w}(p, w) \,.\, X_i(p, w) \quad (11)$$

$$i = 1, \ldots, \ell$$
$$j = 1, \ldots, \ell$$

Da não positividade de K vem que

$$\frac{\partial X_i}{\partial p_j}(p, w) + \frac{\partial X_i}{\partial w}(p, w) \,.\, X_i(p, w) \leq 0\ , \quad i = ,\ldots, \ell$$

Da definição da função utilidade indirecta

$$V(p, w) = U(X(p, w))$$

vem que

$$\psi(p, V(p, w)) = X(p, w)$$

e

$$w = C(p, V(p, w)) = p\, X(p, w) \quad (12)$$

Derivando (12) relativamente a w vem a condição de agregação de Engel

$$1 = \sum_j p_j \frac{\partial X_i}{\partial w} = \frac{\partial C}{\partial V}(p, V(p, w)) \frac{\partial V}{\partial w}(p, w)$$

Derivando (12) relativamente a p_j obtem-se

$$\frac{\partial C}{\partial p_j}(p, V(p, w)) + \frac{\partial C}{\partial V}(p, V(p, w)) \frac{\partial V}{\partial p_j}(p, w) = 0$$

donde

$$-\frac{\frac{\partial V}{\partial p_j}(p,w)}{\frac{\partial V}{\partial w}(p,w)} = \frac{\partial C}{\partial p_j}(p, V(p, w)) = \psi_j\ (p, V(p, w))$$

obtendo-se a identidade de Roy

$$X_j(p, w) = -\frac{\frac{\partial V}{\partial p_j}(p,w)}{\frac{\partial V}{\partial w}(p,w)} \quad (13)$$

De (7), (7'), (10), (13), conclui-se a relação de agregação de Cournot

$$p\, \nabla_p\ x(p, w) = -x(p, w) \quad (14)$$

É claro que as relações obtidas só se dão, pelo menos na forma que foram apresentadas, se se verificarem as condições de diferenciabilidade nas funções nelas envolvidas. Para isso basta impor que a função $U: \mathbb{R}^{\ell} \rightarrow \mathbb{R}$ seja classe C^2, estritamente crescente, estritamente quase-concava.

As relações existentes entre as funções definidas (utilidade, utilidade indirecta, custo, procura Marshalliana, procura Hicksiana) impõem restrições sobre o modelo matemático que se pretende ajustar às observações, permitindo, por outro lado, testar a hipótese, em face dos dados observáveis do consumidor ter regras de comportamento económico como as que o modelo descrito postula.

2.2 Produtor

Supõe-se que o produtor de uma mercadoria dispõe de uma tecnologia de produção representada por uma função

$$F:\ \Delta_1 \times \Delta_2 \rightarrow R$$
$$(x, z) \longmapsto F(x, z)$$

chamada função de produção, onde $\Delta_1 \subset \mathbb{R}_+^{\ell_1}$ é o conjunto dos factores de produção variáveis e $\Delta_2 \subset \mathbb{R}_+^{\ell_2}$ é o conjunto dos factores fixos de produção ($\ell_1 + \ell_2 < \ell$) e que a cada vector (x, z) associa um número real não negativo F(x, z) correspondente à quantidade máxima produtível a partir dos factores de produção (x, z).

Supõe-se que o produtor toma os preços das mercadorias (factores de produção e produto) como dados.

Se F(•, z) for contínua por valores superiores (isto é se $\forall x \in F(\Delta_1, z)$ $L^{\geq}(x, z) = \{x \in \Delta_1 : F(x, z) \geq u\}$ é um fechado) então o problema

$$C(u, p^i, z) = \min_x \{p^i x : F(x, z) \geq u\} \qquad (1)$$

tem solução definindo a função custo C associada à função de produção F(, z) e que goza das seguintes propriedades:

Seja $u \in F(\Delta_1, z)$, $p_i >> 0$, $a > 0$ então

1 - C é não negativa - $C(u, p^i, z) \geq 0$.

2 - C é positivamente homogénea em p
$C(u, \alpha p^i, z) = \alpha C(u, p^i, z)$.

3 - C é não decrescente em p
$p^{i1} \geq p^{i0}$ então $C(u, p^{i1}, z) \geq C(u, p^{i0}, z)$.

4 - C é concava na variável p^i.

5 - C é contínua na variável p^i.

6 - C é não decrescente em u
$u^0 < u^1$ então $C(u^0, p^i, z) < C(u^1, p^i, z)$.

7 - C é contínua por valores inferiores na variável u
$L^-(u, z) = \{v \in F(\Delta_1, z) \mid C(v, p^i, z) \leq C(u, p^i, z)\}$ é um fechado

Se além disso F(•, z) é contínua, crescente (isto é $y >> x \Rightarrow F(y, z) > F(x, z)$) é quase-concava então $F(\Delta_1, z)$, é um intervalo $F(\Delta_1, z) = [\, \bar{u}(z), \bar{\bar{u}}(z)\, [$ onde $0 < \bar{u}(z) \leq \bar{\bar{u}}(z) < \infty$ e C goza das seguintes propriedades adicionais

8 - $C(u, z, p^i)$ é contínua em $p \times F(\Delta_1, z)$

9 - $C(\bar{u}(z), p^i, z) = 0 \qquad \forall p \in P$

10 - $C(u, z, p^i)$ é estritamente crescente na variável u se $p >> 0$

11 - $C(\bar{\bar{u}}(z), p, z) = +\infty$

12 - C(u, pi, z) é estritamente crescente na variável p se $u > \bar{u}(z)$

Estas propriedades são de demonstração imediata decorrendo a maior parte delas das definições.

As observações feitas aquando do estudo do modelo do consumidor, relativamente à hipótese de quase-concavidade da função utilidade são aqui inteiramente válidas havendo apenas que substituir a palavra consumidor por produtor, e a equação preferências (ou utilidade) por conjunto das possibilidades de produção (ou função de produção) - Fig.4.

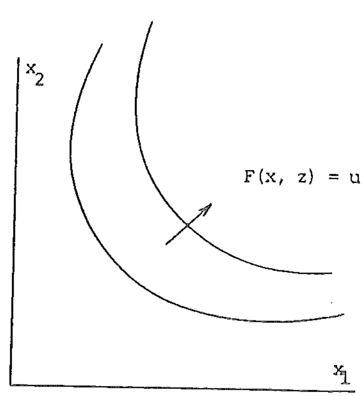


Fig. 4

Suponhamos agora que F é estritamente crescente, estritamente quase concava e de classe C^2. Então o problema (1) tem uma única solução $\psi(u, pi, z)$ ficando definida a função de procura dos factores de produção variáveis.

É claro que todas as considerações feitas relativamente à função custo C e à função de procura ψ, quando F era uma função utilidade, são aqui válidas. Em particular y é contínua, homogénea de grau 0 em p, C é estritamente concava em p, diferenciável, obedecendo ao lema de Shephard

$$\frac{\partial C}{\partial p_j} = \psi_j$$

Se F for fortemente quase-concava, C é de classe C^2 e ψ é diferenciável.

O problema do produtor maximizador do lucro quando da posse do vector **z** de factores de produção fixos é, então, o de determinar o nível de produção **u** solução do problema

$$\max_{u \varepsilon F(\Delta_1, z)} \left[p^0 u - C(u, p^i, z)\right] \tag{2}$$

onde p^0 é o preço do produto e p^i é o preço dos factores de produção variáveis, definido assim uma função lucro,

$$\pi\,(p^0, p^i, z) \tag{3}$$

valor máximo do problema (2) que é positivamente homogénea em (p^0, p^i), crescente em p^0, decrescente em p^i, convexa e de classe C^2 se e só se $C(u, p^i, z)$ for da classe C^2.

As condições de optimalidade conduzem a

$$p^0 = \frac{\partial C}{\partial u}$$

isto é o preço p^0 é igual ao custo marginal de produção.

Fica também definida a função oferta

$$S(p^0, p^i, z)$$

solução do problema e que é homogénea de grau 0 em (p^0, p^i).

Se π for diferenciável em (p^0, p^i) vem que

$$\frac{\partial \pi}{\partial p^0} = S(p^0, p^i, z)$$

$$\frac{\partial \pi}{\partial p^i_j} = - \psi_j(u, p^i, z)$$

onde ψ_j é a procura do factor de produção variável, j.

Se o problema do produtor passar a ser o da aquisição dos factores de produção fixos z e factores de produção variáveis x de modo a produzir a quantidade u, para preços p^i, p^0, maximizar o seu lucro, então o problema global é

$$\max_{z \varepsilon \Delta_2} \left[\pi\,(p^0, p^i, z) - CC_z \times z\right]$$

em que

$$\pi\,(p^0, p^i, z) = p^0\, S(p^0, p^i, z) - C(S(p^0, p^i, z),\; p^i, z)$$

Se S for diferenciável em z, a condição necessária de optimalidade

$$CC_z = \frac{\partial \pi}{\partial z} = p^0\,\frac{\partial C}{\partial S}\,\frac{\partial S}{\partial z} = \left(p^0 - \frac{\partial C}{\partial S}\right)\frac{\partial S}{\partial z}$$

diz-nos que em condições de competividade perfeita, o custo unitário do factor fixo de produção deve igualar a benefício marginal que advem da posse de mais uma unidade desse factor.

2.3 Modelo de escolha discreta

É por vezes impossível conhecerem-se as regras de decisão individuais de cada agente económico duma população homogénea quanto às suas preferências face a um conjunto discreto de alternativas.

Se no processo inferencial se escolherem certos factores como determinantes explícitos na estruturação das preferências dessa população, factores individuais poderão ser introduzidos de forma implícita através de uma função distribuição de probabilidade que indica os desvios de comportamento individual relativamente ao copmortamento médio dessa população.

Neste parágrafo descreve-se o enquadramento teórico de tais modelos de decisão, chamados modelos de escolha discreta. Seguiu-se a primeira parte do trabalho de McFadden sobre o modelo condicional logit.

Seja X um conjunto designado por universo dos objectos de escolha, seja A o conjunto de indivíduos que constituem a população em estudo e seja S um conjunto designado por conjunto de características relevantes para o problema de decisão em causa.

Admitamos que a cada indivíduo $a \varepsilon A$ está associado um vector de atributos $X(a) \varepsilon S$, isto é, existe uma função
e que, além disso, existe uma regra de decisão $h(a, \circ)$ que a cada conjunto finito de alternativas $B \subset X$ associa uma e uma só alternativa $h^*(a, B) \varepsilon B$.

O modelo de escolha individual da população A é o conjunto de regras de decisão individuais

$$H^* = \{h^*(a, \circ) \mid a \varepsilon A\}$$

no qual se introduz uma medida de probabilidade π^*, que induz uma medida de probabilidade π no conjunto

$$H = \{h(s, \circ) \mid s \varepsilon S\}$$

das regras de decisão dos indivíduos com características s.

Suponhamos agora que, entre os indivíduos de A com características s, é tirado à sorte um indivíduo ao qual é apresentado um conjunto finito de alternativas

$$B = \{x_1, \ldots, x_n\}$$

A probabilidade desse indivíduo escolher a alternativa x_i é igual à probabilidade de ocorrência em H duma regra de decisão são $h(s, \circ)$ tal que

$$h(s, B) = x_i$$

isto é

$$P(x_i \mid s, B) = \pi(\{h \varepsilon H \mid h(s, b) = x_i\})$$

Se as regras de decisão dos indivíduos com características $h(s, \circ)$ resultarem da maximização de uma função utilidade (ou de minimização de uma função custo)

$$U(s, \circ) = V(s, \circ) + \xi(s, \circ)$$

onde V é uma função que representa a utilidade "média" dos indivíduos da população e ξ é a componente aleatória então a probabilidade de um indivíduo tirado à sorte da subpopulação com características s, escolher a alternativa x_i, perante o conjinto de alternativas B, é dada por

$$P_i = P(x_i \mid s, B) = \pi(\{h \varepsilon H \mid h(s, B) = x_i\}) =$$

$$= P(\xi(s, x_j) - \xi(s, x_i) < V(s, x_i) - V(s, x_j) \quad \forall_j \neq i \quad i = 1, \ldots, n)$$

A distribuição da probabilidade π induz a distribuição de probabilidade conjunta

$$F(\xi_1, \ldots, \xi_n) = \pi(\{h \varepsilon H \mid \xi(s, x_j) \leq \xi_j, \ j = 1, \ldots, n\})$$

a qual se admitiu derivada parcial relativamente a ξ permite escrever

$$P_i = \int_{-\infty}^{+\infty} F_i(\xi + V_i - V_1, \ldots, \xi + V_i - V_n)\, d\xi \qquad (1)$$

onde

$$V_i = V(s, x_i) \qquad i = 1, \ldots, n$$

isto mostra que dada uma função distribuição conjunta $F(\xi_1, \ldots, \xi_n)$ é possível estimar as probabilidades de selecção de cada alternativa pela expressão (1).

Por exemplo se ξ é tal que as variáveis aleatórias $\xi(s, x_j)$, $j = 1, \ldots, n$ são independentes e identicamente distribuidas com a distribuição de extremos

$$P(\xi(s, x_j) < \xi) = e^{-e^{-\xi}}$$

então as probabilidades de selecção (1) podem escrever-se

$$P(x_i \mid s, b) = \frac{e^{V(s, x_i)}}{\sum_{y \varepsilon B} e^{V(s, y)}} \equiv$$

3. Equilíbrio Económico

Seja M uma economia com M agentes de decisão que produzem,trocam e consomem ℓ mercadorias. Admitamos que as operações de troca se processam em ℓ mercados, um por mercadoria, e seja $P = \mathbb{R}_+^{\ell} - \{0\}$ o espaço do sistema de preços. A cada agente económico i, associamos um vector $w^i \in \mathbb{R}^{\ell}$ de mercadorias, consituindo a sua riqueza inicial, e um conjunto $\Omega_i \subset \mathbb{R}^{\ell}$, o espaço das suas decisões de produção-consumo. Por convenção, se $x^i \in \Omega_i$ for uma decisão de produção-consumo do agente i, então $x_j^i \geq 0$ corresponderá a um consumo de x_j^i unidades de mercadorias j e $x_j^i < 0$ corresponderá a uma produç~ão de $-x_j^i$ unidades de j.

Dado um sistema de preços $p \in P$, uma distribuição de riqueza $w = (w^1, w^2,\ldots, w^M)$ e $\mathbb{R}^{\ell M}$, e as decisões de produção-consumo de outros agentes económicos

$$x^{-i} = \{(x^1,\ldots, x^{i+1}, x^{i+1},\ldots, x^1) \in \Omega_{-1} = \Omega_1 x,\ldots, x\Omega_{i-1}1 \; x\Omega_{i+1}x,\ldots, \Omega_M$$

seja

$$\varphi_i(p, w, x^{-i}) = \{x^i \in \Omega^i \mid px^i \leq pw^i, g_i(x^i, x^{-i}, p, w) < 0, h_i(x^i, x^{-i}, p, w) = 0\}$$

o conjunto das decisões admissíveis, onde g_i e h_i são funções contínuas e $px^i \leq pw^i$ é a restrição orçamental do agente económico i.

Seja

$$f_i : \Omega_i \times \Omega_{-i} \times P \times \mathbb{R}^{\ell M} \to \mathbb{R}$$

a função utilidade do agente económico i, e $u_i(x^{-i}, p, w)$ o conjunto de pontos x_i, solução do problema

$$\max_{z \in \varphi_i(p,w,x^{-i})} f_i(z, x^{-i}, p, w)$$

Dados, p, w e x^{-i} supõe-se que o problema do agente económico i, consiste em escolher uma decisão em $\varphi_i(p, w, x^{-i})$, que maximize a sua utilidade, isto é, escolher um ponto $x^i \in \mu_i(x^{-i}, p, w)$.

Dado $\bar{w}$, um ponto $(\bar{x}, \bar{p}) = (\bar{x}^1,\ldots, \bar{x}^M, \bar{p}) \in \Omega_1,\ldots, \Omega_M \times P$ é um equilíbrio económico para a economia

$$E = (\Omega_i, \varphi_i, f_i)_{i=1,\ldots, M} \qquad \text{sse}$$

1) $\bar{x}^1 \in \mu_i(\bar{x}^{-i}, \bar{p}, \bar{w}), \qquad i = 1,\ldots, M$

2) $\sum_{i=1}^{M} \bar{x}_j^i \leq \sum_{i=1}^{M} \bar{w}_j^i, \qquad j = 1,\ldots, \ell$

Um equilíbrio de Walras para E é um equilíbrio económico onde

$$\sum_{i=1}^{M} \bar{x}^i \leq \sum_{i=1}^{M} \bar{w}^i$$

Considere-se agora um novo agente de decisão, chamado coordenador dos mercados e cujo problema, dadas as decisões $\bar{x}^1,\ldots, \bar{x}^M$ dos outros agentes económicos, consiste em determinar um sistema de preços $\bar{p} \in P$ solução de

$$\max_{p \in \varphi_0(\bar{x},\bar{w})} f_0(p, \bar{x}^1,\ldots, \bar{x}^M, \bar{w}^1,\ldots, \bar{w}^M) \tag{1}$$

onde

$$\varphi_0(\bar{x}, \bar{w}) = P = \Omega_0 \qquad \forall(\bar{x}, \bar{w})$$

e

$$f_0(p, \bar{x}, \bar{w}) = p \times \left[\sum_{i=1}^{M} \bar{x}^1 - \sum_{i=1}^{M} \bar{w}^i\right]$$

Pelo teorema de Arrow-Debreu-Nash (ver anexo) pode-se provar a existência e unicidade de um equilíbrio de Nash para o sistema $N = (\Omega_i, \varphi_i, f_i)$ $i = 0,\ldots, M$ e consequentemente um equilíbrio económico para a economia $E = (\Omega_i, \varphi_i, f_i)$ $i = 1,\ldots, M$.

Observação 1

Se $(\bar{x}, \bar{w})$ é um equilíbrio económico para o sistema $(\Omega_i, \varphi_i, f_i)$ $i = 1,\ldots, M$ e para algum $\sum_{i=j}^{M} \bar{x}_j^i < \sum_{i=1}^{M} \bar{w}_j^i$ então $\bar{p}_j = 0$, pois de contrário $\bar{p}$ não seria ma solução do problema (1); se $\bar{p}_j > 0$ vem necessariamente que $\sum_{i=1}^{M} \bar{x}_j^i - \sum_{i=1}^{M} \bar{w}_j^i = 0$ o que é equivalente à relação de complementaridade.

$$\bar{p}_j \left(\sum_{i=1}^{M} \bar{x}_j^i - \sum_{i=1}^{M} \bar{w}_j^i \right) = 0 \qquad , j = 1,\ldots, \ell$$

que se deve verificar num ponto de equilíbrio e que na literatura económica é conhecida por *identidade de Walras.*

Observação 2

Se as funções f_i e as correspondentes φ_i só dependerem do sistema de preços p, as correspondências μ_i também só dependem do sistema de preços; se além disso para cada sistema de preços, $\mu_i(p)$ e um conjunto com um só elemento, isto é, μ_i é uma função dependente do sistema de preços, então o problema de se encontrar um equilíbrio $(\bar{x}, p)$ é equivalente a encontrar uma solução $\bar{p}$ do problema de complementaridade

$$g(p) \geq 0$$
$$p\, g(p) = 0$$
$$p > 0$$

onde

$$g(p) = \sum_{i=1}^{M} \left(w^i - \mu_i(p) \right)$$

4. Descrição do Modelo DFI

As actividades que constituem o sector energético, desde a exploração de recursos naturais e importação de combustíveis ou matérias primas até às que caracterizam o consumo final de energia, situam-se nos nós dum grafo orientado, sem circuitos, cujos arcos representam a informação útil (quantidades e preço) entre as actividades, permitindo assim, a sua interligação.

A cada produtor homogéneo está associado um mercado específico que permite modelar o comportamento dos consumidores, quer sejam produtos de uma forma de energia, quer sejam consumidores de energia final.

Dada a procura global de energia correspondente a um cenário de desenvolvimento da economia, o modelo DFI baseado na teoria do equilíbrio económico geral, determina a expansão óptima do sistema energético de modo a satisfazer a procura global de energia correspondente àquele cenário, indicando como irão competir as diversas tecnologias de produção na satisfação daquela procura, tendo em conta as preferências individuais, a incerteza característica de cada um dos mercados, e condicionamentos de ordem tecnológica a que os processos de transformação estão sujeitos.

A sua implementação é conseguida através dos seguintes elementos básicos:

- *processos* que descrevem as diferentes actividades (produtor e consumidor).
- *rede* (grafo orientado) que interliga os processos, possibilitando a sua interacção.
- *algoritmo* que conduz à determinação dos valores numéricos das variáveis do modelo.

4.1 O mercado da energia

O objectivo do mercado é a selecção das quantidades associadas às diferentes tecnologias concorrentes ao seu abastecimento de modo a minimzar o custo global da satisfação da procura.

No sector energético o mesmo bem "energia" é produzido por diferentes processos tecnológicos com custos e características perfeitamente conhecidos. No entanto aquela que é apropriada (de mais baixo custo de produção) para um produtor poderá não ser para o próximo. Assim nasce uma incerteza, por parte do mercado, relativamente aos preços oferecidos pelos diferentes produtores concorrentes ao seu abastecimento. É neste contexto de incerteza que se formulará o problema do consumidor.

Cada produtor representa um conjunto de processos semelhantes que utilizam o mesmo tipo de tecnologia (por ex. a produção hidroeléctrica deve representar todo um conjunto de centrais hídricas...) existindo, então, uma função densidade de probabilidade dos preços de oferta para aquele produtor.

Admitindo que a incerteza do preço de oferta p_i, associada à tecnologia i de um conjunto de produtores, tem uma distribuição de Weibull

$$f_i(p_i) = \left(\frac{\gamma_i}{\beta_i} \right) \times \left(\frac{p_i}{\beta_i} \right)^{\gamma_i - 1} \times e\left(\frac{-p_i}{\beta_i} \right)^{\gamma_i} \qquad (8)$$

com

$\gamma_i > 0;\ \beta_i > 0;\ p_i > 0$

γ_i - parâmetro característico da forma da distribuição.

β_i - parâmetro de escala.

então, a esperança matemática do preço ser igual a p_i é

$$E_i(p_i) = \bar{p} = \int_0^\infty p_i\, f_i(p_i)\, dp_i = \beta_i\, \Upsilon(1 + \frac{1}{\gamma_i}) \qquad (9)$$

com

$$\Upsilon(\gamma_i) = \int_0^\infty p_i^{\gamma_i - 1}\, e^{p_i}\, dp_i$$

O tipo de tecnologia representativa de uma actividade genérica (ex.: produção de energia mecânica através de motores de combustão) abrange um mesmo tipo de produtores (ex.: a gasolina ou diesel) pelo que podemos admitir que a forma das curvas das respectivas funções densidade de probabilidade é a mesma. Significa isto que para um determinado mercado (ex.: mercado automóvel) o parâmetro de forma γ_i é o mesmo para todas as tecnologias concorrentes, ou seja,

$$F_i(p_i) = 1 - e - \left(\frac{p_i}{\beta_i} \right)^{\gamma} = \int_0^\infty f_i(p_i)\, dp_i \qquad (10)$$

$$\gamma_i = \gamma\ ,\ \forall_i$$

O nosso problema consiste em determinar a fracção do mercado afecta a determinada tecnologia concorrente a esse mercado. Essa fracção será, num primeiro passo, igual à probabilidade do preço

$$s_i = \text{Prob}\ .(p_i < p_j),\ \forall_j \neq i \qquad (11)$$

de oferta da tecnologia i ser inferior a todos os preços associados às restantes tecnologias (a figura dá-nos um exemplo para o caso de duas tecnologias).

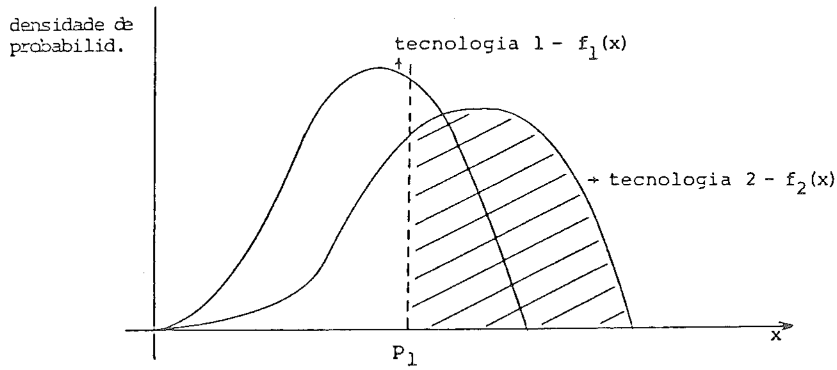


Fig. 5

$$s_1 = \text{Prob. } (p_1 < p_2) = \int_{p_1}^{\infty} f_2(x)\, dx$$

O segundo passo, corresponde à ponderação desta probabilidade da tecnologia i ter um preço de oferta igual a p_i (p_1 na figura), dum modo geral

$$s_i = \int_0^{\infty} f_i(p_i) \times \left[\prod_{\substack{i=1 \\ i \neq j}}^{N} \int_{p_i}^{\infty} f_j(p_j) \right] dp_i \qquad (12)$$

s_i - fracção de mercado afecta à tecnologia i.
N - número de tecnologias concorrentes ao mercado.

Atendendo às expressões anteriores ((8), (9) e (10))

$$s_i = \frac{\bar{p}_i^{\,-\gamma}}{\sum_{j=1}^{N} \bar{p}_j^{\,-\gamma}}$$

e se tivermos em consideração as preferências do consumidor relativamente a cada um dos produtores, em igualdade de preços de oferta, teremos a expressão da fracção de mercado que cabe ao produtor i, concorrente com os outros N-1

$$s_i = \frac{f_i\, \bar{p}_i^{\,-\gamma}}{\sum_{j=1}^{N} f_j\, \bar{p}_j^{\,-\gamma}} \qquad (14)$$

Num processo de mercado existe ainda um factor importante a considerar: a inércia, por parte dos consumidores, a responder às variações relativas dos preços de oferta das tecnologias concorrentes ao longo de um período. Daqui se infere da necessidade de ponderação ao longo da fracção de mercado $s_i(t)$, por um factor k

$$\bar{s}_i(t) = k \times s_i(t) + (1 - k)\, \bar{s}_i\, (t - 1) \qquad (15)$$

k - factor de ponderação que traduz a maior ou menor rapidez de ajuste do consumidor a variações relativas dos preços de oferta. -Fig. 7

$\bar{s}_i(t)$ - fracção do mercado em t, afecta ao produtor i.

$s_i(t)$ - fracção do mercado t, admitindo a não existência de atraso na resposta (equação 14).

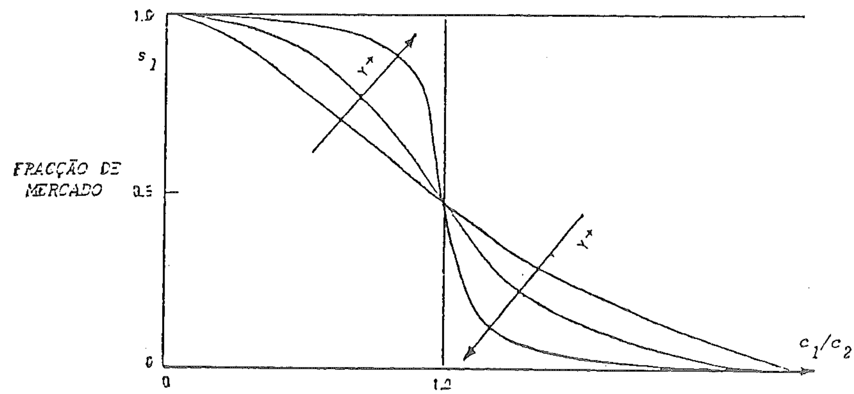


Fig. 6

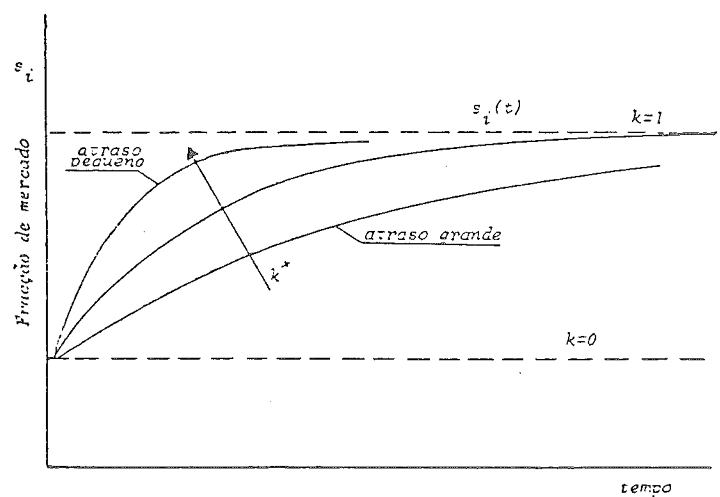


Fig. 7

A finalizar:

- O preço que se estabelece no mercado, será igual à média dos preços de
- A quantidade total de energia oferecida pelos produtores terá de ser igual à quantidade procurada pelos consumidores, isto é, a oferta deve igualar a procura.

4.2 O produtor no mercado energético

No modelo D.F.I., à actividade correspondente a uma determinada tecnologia, associamos um produtor do bem energia que necessita de L factores de produção para as capacidades instaladas new(t) em t=1,..., T.

Seja ainda:

- $p(\tau)$, o preço do produto de venda em τ.
- $CC(\tau)$, o custo unitário da capacidade instalada em τ.
- $i=1, 2, \dots, L$, o conjunto dos factores necessários à produção de energia.
- $f_i(t,\tau)$, o multiplicador que traduz o efeito do tempo sobre os factores de

produção a_i que são necessários no início da exploração.

$$a_i(t,\tau) = a_i\, f_i(t,\tau)$$

- $cf(t,\tau)$, o factor de utilização em τ, da capacidade instalada em t

$$cf(t\,\tau) = \frac{\beta}{1 + (\alpha \dfrac{\varnothing(t,\tau)}{p(\tau)})^{\gamma}}$$

sendo α, β, γ, os parâmetros associados à tecnologia utilizada.

O encargo variável para a produção de uma unidade do bem energia, relativo à capacidade instalada em t é

$$\varnothing(t,\tau) = \sum_{i=1}^{L} a_i(t,\tau)\, p_i(\tau)$$

sendo $p_i(t)$ o preço em t dos factores de produção i.

Deste modo, tendo em vista a maximização do lucro, o problema do produtor é

$$\max_{new(\tau)} \sum_{\tau=1}^{T} \left\{ p(t) \times \sum_{t=t_0}^{\tau} new(t)\, cf(t,\tau) - \left[\sum_{t=t_0}^{\tau} new(t)\, cf(t,\tau)\, \varnothing(t,\tau) \right] - cc(\tau)\, new(\tau) \right\}$$

com t_0 correspondente ao início da exploração.

A condição necessária de optimalidade, conduz-nos a

$$cc(\tau) = \sum_{t=\tau}^{T} (p(t) - \varnothing(\tau,t)) \times cf(\tau,t)$$

Multiplicando ambos os membros da equação anterior por $new(\tau)$ a equação resultante traduz o seguinte:

> O investimento efectuado no período t, deverá ser compensado pela soma actualizada dos benefícios futuros dele decorrentes.

De notar que da expressão se poderá obter uma relação recorrente no tempo do preço de venda do produto ("energia").

$$p(t) = \frac{cc(\tau)}{cf(\tau,\tau)} + \varnothing(\tau,\tau) - \sum_{t=\tau+1}^{T} \frac{(p(t) - \varnothing(\tau,t))}{cf(\tau,\tau)}\, cf(t,\tau)$$

Como já foi anteriormente demonstrado, uma vez fixa a capacidade de produção, o preço de venda deverá igualar o custo marginal de produção, ou seja,

$$\lambda(\tau) = p(\tau) \equiv \frac{\partial C(q,p^i,p^\circ)}{\partial q}$$

4.3 Algoritmo

O algoritmo que o modelo utiliza para determinação do sistema de preços p* de equilíbrio, é baseado no teorema de ponto fixo, utilizando a estrutura particular do sector energético em estudo.

Como vimos, as actividades que constituem o sector energético estão representadas

Consideremos, então, o grafo representativo da cadeia das actividades do sector energético, orientado no sentido da importação (ou extracção de recursos) para o consumo final, e seja:

$\Gamma^+(x)$ - o conjunto de sucessores do nó x.

$\Gamma^-(x)$ - o conjunto de antecessores do nó x.

$r(x)$ - nível do nó x.

Descrição geral

1. Para cada nó x do grafo são determinados os conjuntos $\Gamma^+(x)$, $\Gamma^-(x)$ e o seu nível r(x).
2. Percurso ascendente do grafo - "UP".
   Conhecidas as quantidades a produzir (oferta) por cada uma das tecnologias consideradas, determina, sequancialmente o sistema de preços p - [p = U(Q)].
   2.1 Nível n=1
   Nos nós do primeiro nível como os das importações e os dos recursos, os
   2.2 Nível n
   Para cada actividade correspondente ao nó x de nível r(x)=n, resolve o problema de optimização, determinando os preços p_x de oferta do bem energia, tendo em conta os preços p_y das actividades dos níveis inferiores, isto é, y ε $\Gamma^-(x)$.
   2.3 Nível n+1 (sobe de nível)
   Se n < N segue para 2.2. Caso contrário, vai para 3.
3. Percurso descendente do grafo - "DOWN"
   Dado um sistema de preços p, determina sequencialmente as quantidades afectas a cada uma das tecnoloigas concorrentes a um dado mercado [Q=D(p)].
   3.1 Nível N
   Neste nível, os nós possuem já as procuras definidas exogenamente.
   3.2 Nível n
   Para cada uma das actividades x do nível r(x)=n, resolve o problema de optimização associado, satisfazendo a procura definida em todas as actividades y de nível superior, isto é, a procura dos nós y ε $\Gamma^+(x)$.
   3.3 Nível n-1 (desce de nível)
   Se n > 0 segue para 3.2. Caso contrário, segue para 2, iniciando uma nova iteração, caso não tenha atingido ou a convergência, ou o número máximo de iterações.

Vimos que a uma iteração correspondem dois percursos no grafo: um, no sentido ascendente, e outro, no sentido descendente. Enquanto que na fase ascendente o problema a resolver é encontrar p face às quantidades Q — p=U(Q) —, na fase descendente, o problema resume-se a determinar as quantidades Q face ao sistema de preços p — Q = D(p) — .

O problema de encontrar a solução p* que satisfaça as relações de equilíbrio, resume-

$$p = U(Q) = U[D(p)] = f(p)$$
$$Q = D(p) = D[U(Q)] = g(Q)$$

ANEXO

Teorema de Arrow-Debreu-Nash

Seja $N = (\Omega_i, \varphi_i, f_i)$ $i=1,\ldots, N$ um sistema de N centros de decisão, onde para o centro de decisão i

$\Omega_i \subset \mathbb{R}^{N_i}$ conjunto maximal das decisões admissíveis

$\Omega_{-i} = \Omega_1 \times \ldots \times \Omega_{i-1} \times \Omega_{i+1} \times \Omega_N$

$\varphi_i : \Omega_{-i} \to \Omega_i$ correspondências que a cada vector de decisão $x^{-i} = (x^1,\ldots, x^{i-1}, x^{i+1},\ldots, x^N)$ dos outros centros de decisão associa o conjunto $\varphi_i(x^{-i}) \subset \Omega_i$ das decisões admissíveis

$f_i : \Omega_i \times \Omega_{-i} \to \mathbb{R}$ função utilidade

$$\mu_i(x^{-i}) = \{x^i \in \varphi_i(x^{-i}) : f(x^i, x^{-i}) = \max_{z \in \varphi_i(x^{-i})} f(z, x^{-i})\}$$

Se para cada $i \in \{1,\ldots, N\}$, Ω_i for um conjunto convexo e compacto, φ_i uma correspondência contínua com imagens convexas não vazias e f_i uma função contínua e quase concava na primeira variável, então, existe um equilíbrio de Nash para o sistema N, isto é, existe um ponto $\bar{x} = (\bar{x}^1,\ldots, \bar{x}^N)$ tal que

$$\bar{x}^i \in \mu_i(\bar{x}^{-i}), \quad i=1,\ldots, N$$

Bibliografia

1 - *Théorie de L'équilibre Général et Économique du Bien-être*
- J. Quik e R. Saposnik - Presses Univ. de France (1974)

2 - *Leçons de Théorie Microeconomique*
- E. Malinvaud - Bordas (1975)

3 - *Microeconomic Theory. A Mathematical Approach*
- R. Quandt e J. Henderson - McGraw Hill (1980)

4 - *Descrição do Modelo de Gestão de Base de Dados do DFI*
- Victor Baptista e M. Natália Tavares - DT 110/82/OCPL

5 - *Apresentação do Módulo de Equilíbrio Económico (GEMS)*
- Victor Baptista e M. Natália Tavares - DT 126/83/OCPL

6 - *Apresentação das Alterações Introduzidas ao Modelo DFI*
- Victor Baptista e M. Natália Tavares - DT 127/83/OCPL

7 - *Generalized Equilibrium Modeling: The Methodology of the SRI - Gulf Energy Model*
- Ed. Cazalet - Decision Focus Incorporated

8 - Frontiers in Econometrics
- Academic Press, New York (1973)

Investigação Operacional, Vol.7 nº1 (Junho/1987) 109 - 119

METODOLOGIA NA MODELAÇÃO DO CONSUMO ENERGÉTICO A NÍVEL DISTRITAL - ALGUNS RESULTADOS -

J.A. Assis Lopes

Secção de Urbanização e Sistemas
Instituto Superior Técnico

Resumo: É bem representativa a fracção do consumo energético nacional que está associada ao sector dos transportes (cerca de 27% em termos de energia final) o qual absorve quase metade do consumo nacional dos derivados líquidos de petróleo (gásoleo e gasolina) pelo que é de manifesto interesse procurar modelar o comportamento do consumidor nacional, neste sector, o que, poderá conduzir a políticas mais realistas que contribuam para uma poupança energética neste sistema.

Num primeiro trabalho já realizado obtiveram-se alguns resultados importantes nomeadamente:

- explicação do consumo nacional de gasolina em TI.
- fluxos em transporte colectivo de passageiros (ferroviário urbanos e interurbanos)
- custos e previsões energéticas no sector.etc.

Tornou-se relevante prosseguir o estudo de forma a modelar o consumo inter e intra distrital, factor explicativo particularmente importante do grau de abertura ou isolamento em relação às trocas efectuadas no domínio comercial e cultural, no estreitamento das relações inter-regionais e da mobilidade nas restantes actividades que caracterizam a sociedade em si.

Apresentamos seguidamente uma primeira abordagem na formulação do problema e os resultados obtidos.

I - Metodologia

Tendo em vista os objectivos atrás referidos e seguindo ordem idêntica àquela por nós apresentada em "Modelação Econométrica para o Consumo Energético no Sector de Transportes" (I Congresso-Nacional da APDIO - 1982) analisámos prioritariamente o transporte individual e neste contexto o consumo correspondente a gasolina já que o devido ao gásoleo, comparado com o que é empregue nos transportes públicos e de mercadorias representa uma pequena parcela apenas.

Era de fundamental importância conseguir por um lado explicar os gastos em gasolina (normal + super) a nível distrital e por outro lado obter uma amostra dos fluxos (origem/destino) onde fossem especificados os modos de transporte utilizados (individual, mercadorias, público, etc.). Poder-se-ia então integrar esta informação numa formulação do tipo da apresentada na figura 1, a qual, permitiria, prever o consumo sectorial de combustíveis líquidos a nível de distrito e analisar a sensibilidade deste a variações dos factores explicativos, nomeadamente de natureza económica.

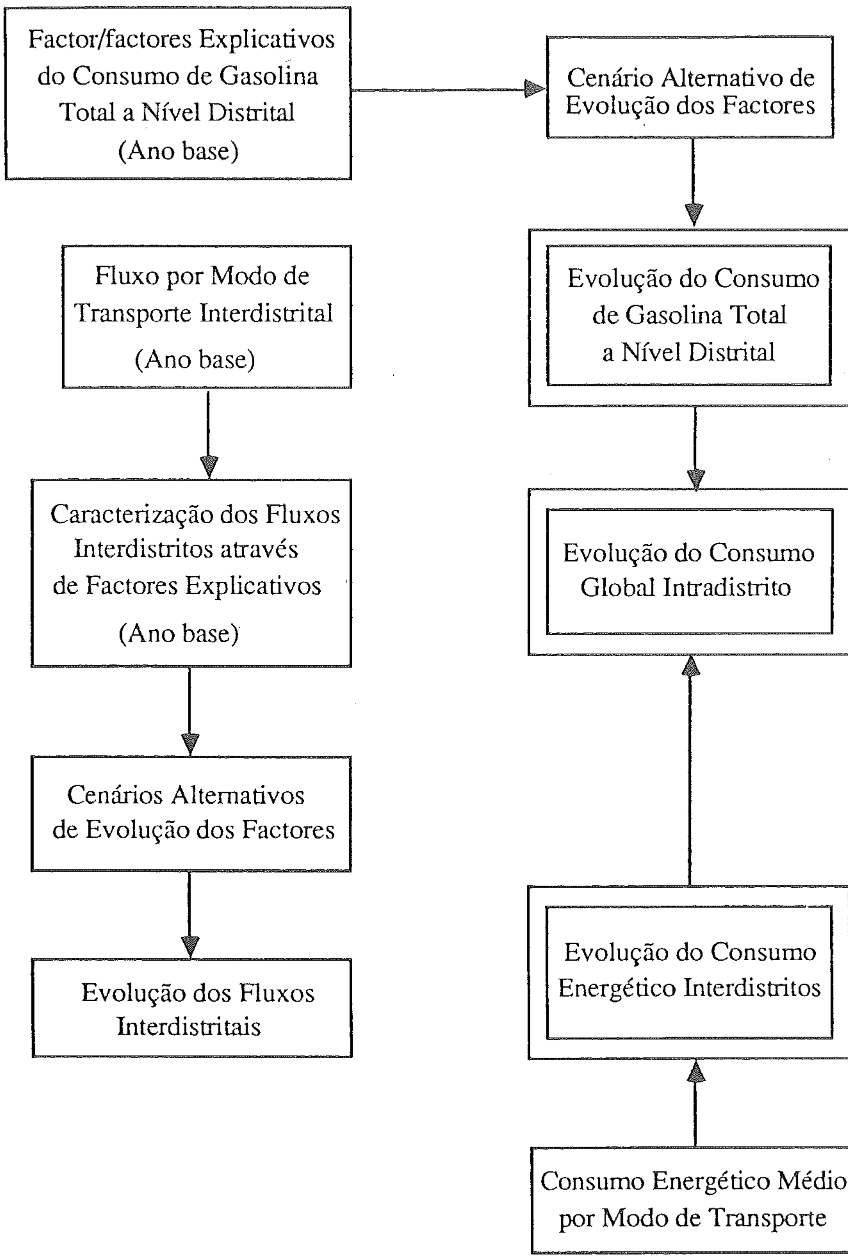


Fig.1 Modelação do Consumo Energético a Nível Distrital

II - Modelação do Consumo de Gasolina a Nível Distrital (TI):

Tendo em vista conseguir uma explicação do consumo de gasolina a nível distrital efectuou-se uma pesquiza exaustiva dos indicadores estatísticos existentes, tendo em conta que:

a. Haveria uma relação privilegiada com indicadores de natureza económica.
b. Verificar-se-ia naturalmente uma relação directa com a riqueza criada no distrito.
c. O grau de ruralidade ou urbanidade não seriam factores estranhos à sua explicação nomeadamente por motivos de natureza social, etc.

Assim, seleccionámos o Valor Acrescentado Bruto (VAB) distrital para a indústria transformadora e a densidade demográfica distritais como indicadores que, disponíveis, melhor se adaptam ao fim em vista, obtendo-se para o ano base (1979) o modelo seguinte, o qual, corrobora as hipóteses anteriormente formuladas:

$$C\,G^{i}_{1979} = 4984.826 + 1.918 \quad VAB^{i}_{1979} + 165.730 \;\; DP^{i}_{1979} \qquad r^{2}_{O.P.} = .927$$

onde:

CG^{i}_{1979} = Consumo de gasolina (M^3) para o distrito i ano 1979.

DP^{i}_{1979} = Densidade populacional (Hab/Km^2) no distrito i ano 1979.

VAB^{i}_{1979} = Valor acrescentado bruto (x10^6 esc.) no distrito i no ano 1979.

Apresentam-se no quadro I e figura 2 os valores observados e previstos para o ano base.

QUADRO I

VALORES OBSERVADOS E PREVISTOS DO CONSUMO DE GASOLINA (M^3) DISTRITAL PARA O ANO BASE

DISTRITO	V.OBSERVADO*	V. PREVISTO MODELO
AVEIRO	70017	80635
BEJA	15118	8930
BRAGA	53075	86262
BRAGANÇA	10068	9994
C. BRANCO	18512	17519
COIMBRA	44448	38257
EVORA	18166	12268
FARO	50517	20429
GUARDA	16849	15106
LEIRIA	49187	42710
LISBOA	278651	232427
PORTALEGRE	13048	11892
PORTO	159660	198642
SANTAREM	46692	29804
SETUBAL	75630	79517
V. CASTELO	19230	28786
V. REAL	15995	16526
VISEU	27068	22138

* Fonte DGE

Utilizando agora o modelo atrás especificado e admitindo que:

a. A taxa média distrital de variação demográfica anual é idêntica á média verificada entre os censos de 1970 e 1981. (1.49%-INE)

b. O crescimento do VAB distrital acompanhou para os anos de 1981, 82 e restantes a variação do PBI.

c. As variações do PBI foram de acordo com as taxas apresentadas no Quadro II.

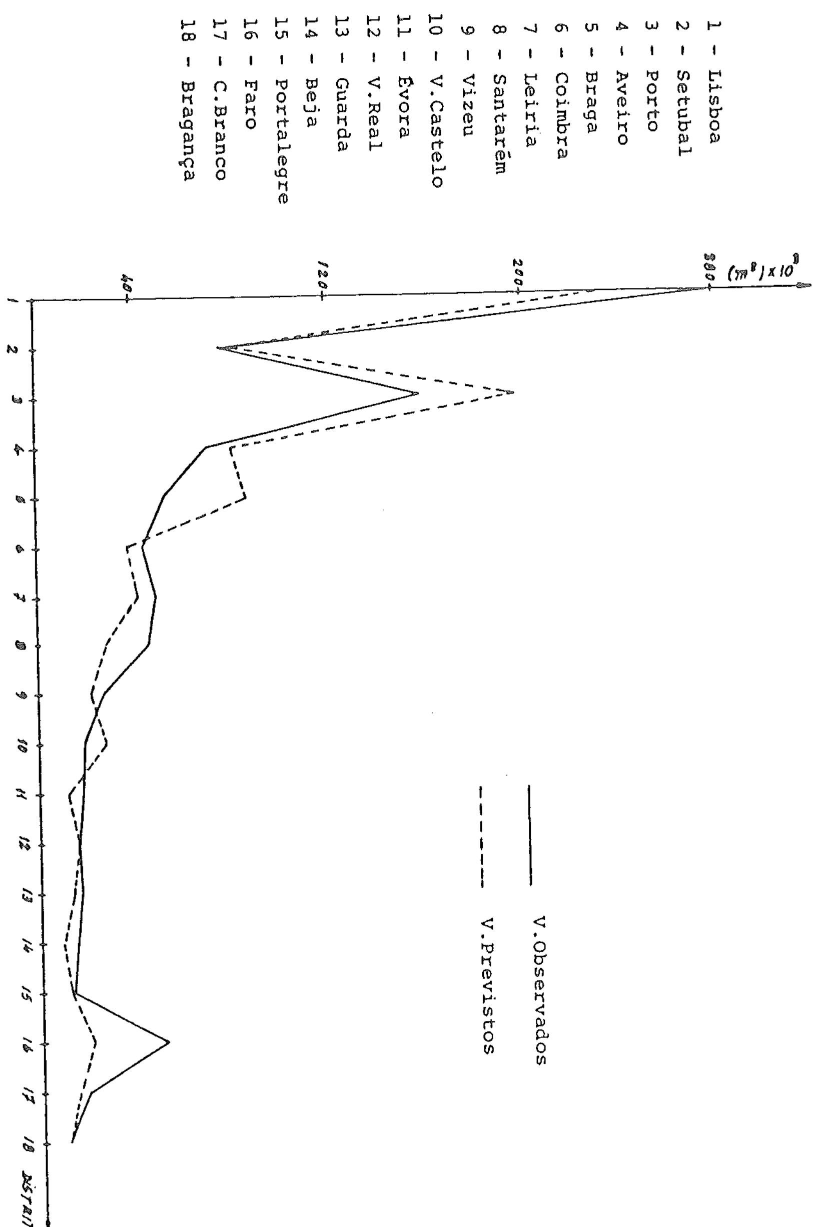


Fig. 2 - Valores observados e previstos do consumo de gasolina a nível distrital (ano base)

QUADRO II

VARIAÇÃO DO PBI AOS PREÇOS DE MERCADO

ANOS	% VARIAÇAO
1980 - 81	.80(1)
1981 - 82	3(1)
1982 - 83	.80(2)
1983 - 84	-1.4(2)

(1) Relatório do C. ADM B. Portugal
(2) Valores Previstos pelo IACEP

Obtemos as seguintes previsões para o consumo de gasolina em 1984 (QUADRO III):

QUADRO III

CONSUMO DE GASOLINA (m^3) PREVISTO PARA 1984 A NÍVEL DISTRITAL

DISTRITO	1984
AVEIRO	102439
BEJA	8962
BRAGA	101154
BRAGANÇA	10218
C. BRANCO	20530
COIMBRA	43306
EVORA	13309
FARO	21640
GUARDA	15668
LEIRIA	51597
LISBOA	270479
PORTALEGRE	12501
PORTO	237038
SANTAREM	33663
SETUBAL	106782
V. CASTELO	32028
V. REAL	17295
VISEU	24489

III - Modelação do Fluxo Interdistrital:

Com o objectivo de modelar o fluxo interdistrital utilizou-se a matriz origem/destino construida através das contagens anuais de viagens interdistritais (por modo de transporte e motivação de viagem) cuja amostra foi levada a cabo pela JAE em 1979, sendo representado na figura 3 as suas principais linhas de força e no Quadro IV os resultados médios do referido estudo, para o número total de viagens tendo o distrito como origem ou destino.

De igual modo usou-se o valor Acrescentado Bruto (VAB) distrital para a indústria transformadora como o indicador que melhor se adaptava ao fim em vista.

Vários modelos foram ensaiados tendo sido os mais significativos aqueles que relacionavam o fluxo médio com a percentagem da contribuição do VAB distrital para o total continental, conforme se especificam:

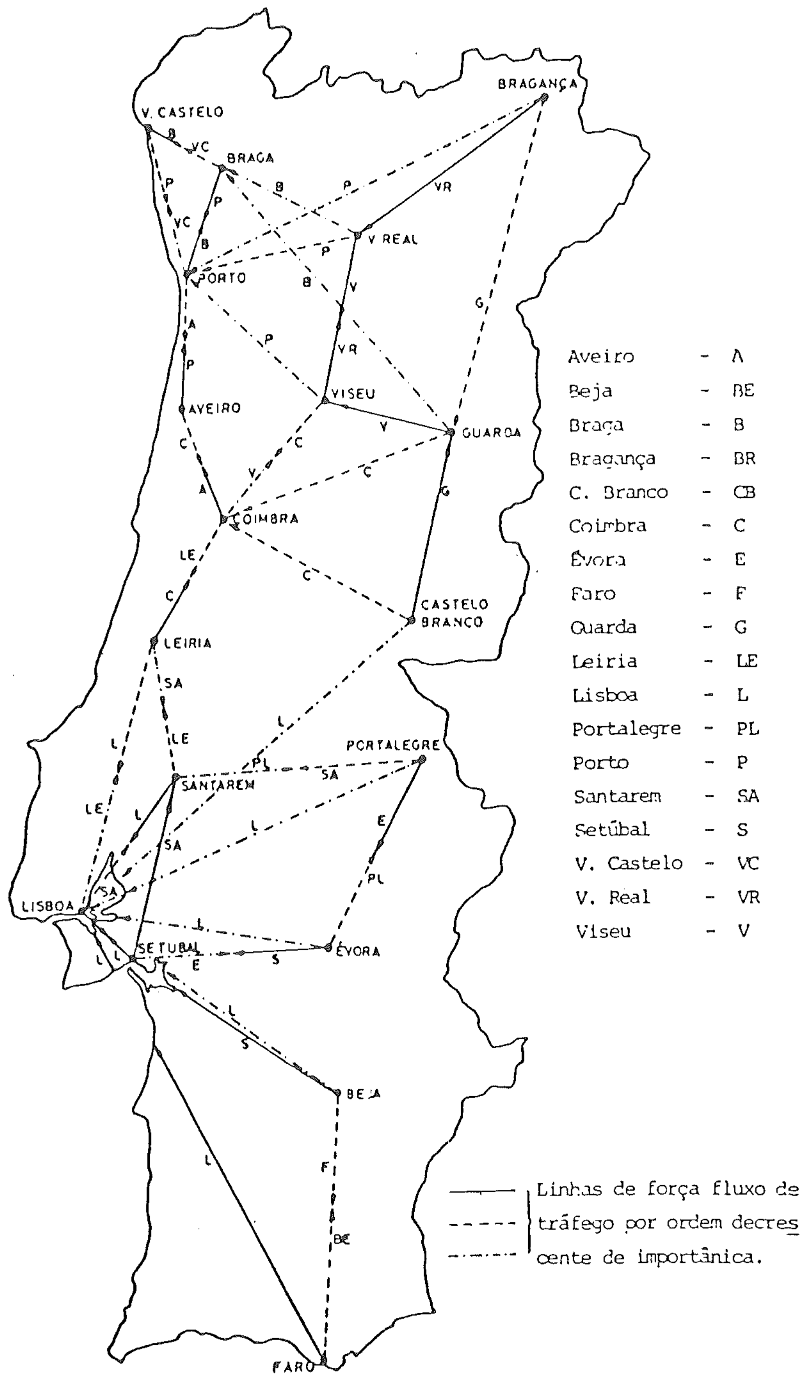

BRAGANÇA
V. CASTELO
BRAGA
V. REAL
PORTO
AVEIRO
VISEU
GUARDA
COIMBRA
CASTELO BRANCO
LEIRIA
PORTALEGRE
SANTAREM
LISBOA
SETUBAL
ÉVORA
BEJA
FARO
Aveiro - A
Beja - BE
Braga - B
Bragança - BR
C. Branco - CB
Coimbra - C
Évora - E
Faro - F
Guarda - G
Leiria - LE
Lisboa - L
Portalegre - PL
Porto - P
Santarem - SA
Setúbal - S
V. Castelo - VC
V. Real - VR
Viseu - V
Linhas de força fluxo de tráfego por ordem decrescente de importânica.

Fig. 3

QUADRO IV

NÚMERO MÉDIO DE VIAGENS INTER-DISTRITAIS

($x10^3$)

DISTRITO	Nº MÉDIO DE VIAGENS TENDO O DISTRITO COMO ORIGEM OU DESTINO NO ANO BASE (1979) $x10^3$
AVEIRO	3615
BEJA	482
BRAGA	3397
BRAGANÇA	212
C. BRANCO	365
COIMBRA	2413
ÉVORA	724
FARO	324
GUARDA	503
LEIRIA	2151
LISBOA	8402
PORTALEGRE	357
PORTO	7070
SANTAREM	2098
SETUBAL	7374
V. CASTELO	853
V. REAL	585
VISEU	1002

(A) $FL_i = 489.2 + 3.3132 \times 104 \quad PVAB_i \qquad r^2 = .930$

(B) $LG_e\, FL_i = 9.755 + .6784 \quad LG_e\, PVAB_i \qquad r^2 = .840$

onde

FL_i = Fluxo médio ($x10^3$ viagens) tendo o distrito i como origem ou destino.

$PVAB_i$ - % do VAB do distrito i em relação ao VAB total continental para a indústria transformadora.

Apresentam-se no Quadro V e figura 4, os resultados obtidos.

QUADRO V

VALORES OBSERVADOS E PREVISTOS DO Nº MÉDIO DE VIAGENS TENDO O DISTRITO COMO ORIGEM OU DESTINO PARA O ANO 1979 ($x10^3$)

DISTRITO	Nº MÉDIO DE VIAGENS OBS.	Nº MÉDIO DE VIAG. PRE.P/ MOD. (A)	Nº MÉDIO DE VIAG. PRE.P/ MOD. (B)
AVEIRO	3615	3703	3584
BEJA	482	555	263
BRAGA	3397	3636	3534
BRAGANÇA	212	522	164
C. BRANCO	365	1052	1109
COIMBRA	2413	1715	1873
EVORA	724	754	667
FARO	324	853	827
GUARDA	503	820	776
LEIRIA	2151	1980	2137
LISBOA	8402	9136	6980
PORTALEGRE	357	721	610
PORTO	7070	7678	6164
SANTAREM	2098	1615	1769
SETUBAL	7374	4928	4455
V. CASTELO	853	920	926
V. REAL	585	588	345
VISEU	1002	754	667

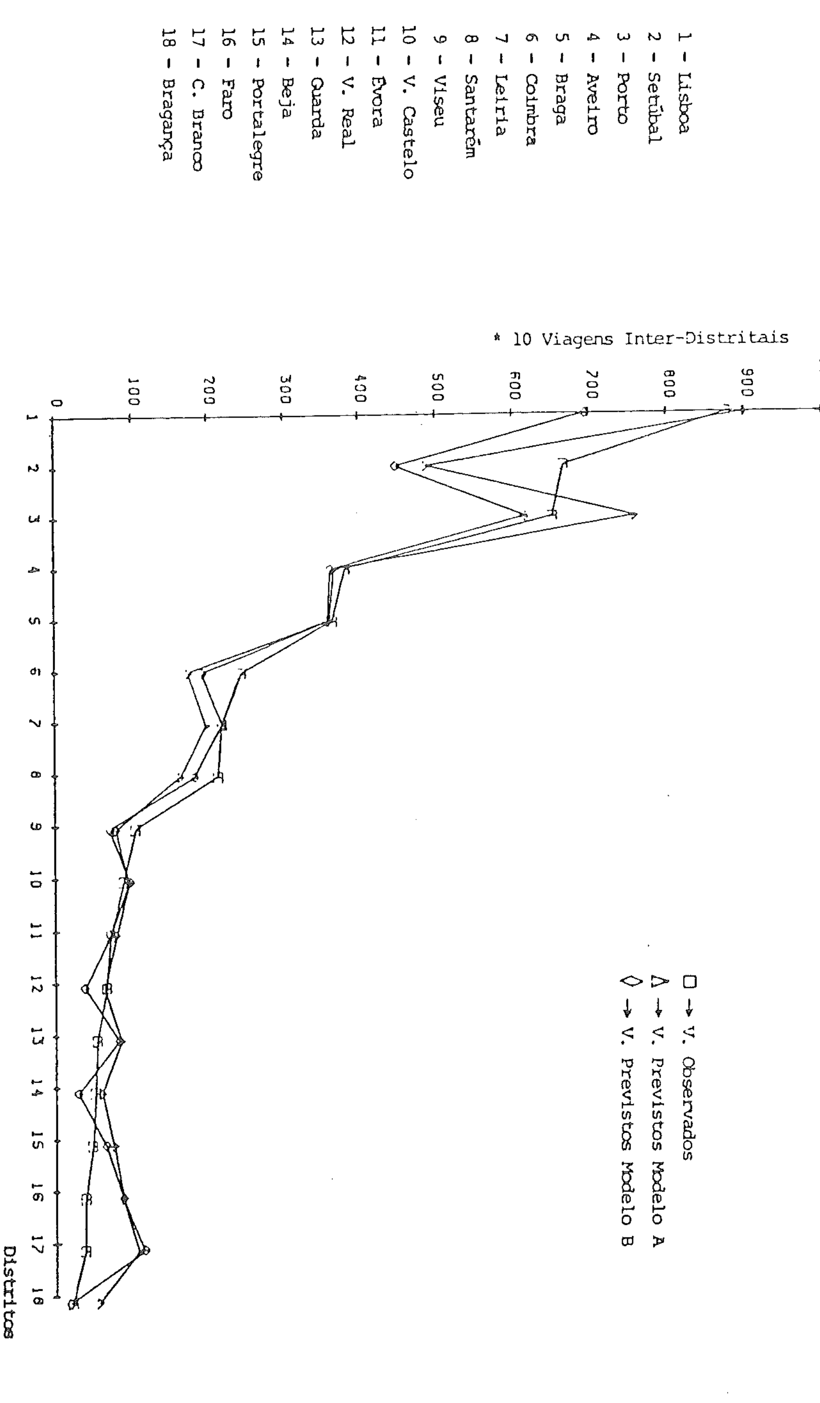


Fig. 4 - Valores Observados e Previstos dos Fluxos Inter-Distritais (Modelo A e B)

Se por hipótese admitirmos que:

a. As condições atrás referidas se mantêm no tocante à variação do VAB.

b. O consumo de gasolina do automóvel privado aos 100km é aproximadamente de 7.5 litros.

c. A autonomia do automóvel privado em capacidade de percurso situa-se próxima dos 300km. (valor médio)

d. A distância média percorrida por viagem em 1979 (ano base) é mantida variando apenas o seu número.

será possível prever as viagens médias efectuadas tendo o distrito como origem ou destino e o consumo médio em combustível que lhe está associado, conforme se representa no Quadro VI.

QUADRO VI

VIAGENS MÉDIAS INERDISTRITAIS (x10^3) E CONSUMO MÉDIO (m^3)
PREVISTOS PARA 1984

Distrito	Distância Média por viagem tendo o distrito como fonte abastecedora (Km)	Viagens Médias tendo o distrito como origem ou destino (10^3)	Consumo Médio em viagens interdistritais (m^3)
		1984	1984
Aveiro	106.6	4008	32044
Beja	214.7	539	8679
Braga	112.9	3557	30119
Bragança	261.3	516	10112
C. Branco	225.7	1069	18096
Coimbra	144.9	1632	17736
Evora	177.2	721	9582
Faro	249.9	814	15256
Guarda	195.9	744	10931
Leiria	181.6	2043	27826
Lisboa	147.3	8709	96213
Portalegre	222.3	705	11754
Porto	137.6	7732	79794
Santarém	166.3	1510	18833
Setúbal	114.4	5293	45413
V. Castelo	123.7	943	8749
V. Real	232.4	592	10319
Viseu	213.9	787	12625

Tornando-se obvia a determinação, por diferença, da gasolina em viagens interdistritais, conforme abaixo especificada (QUADRO VII) e se apresenta na figura 5.

1 - Lisboa
2 - Setúbal
3 - Porto
4 - Aveiro
5 - Braga
6 - Coimbra
7 - Leiria
8 - Santarém
9 - Viseu
10 - V.Castelo
11 - Évora
12 - V.Real
13 - Guarda
14 - Beja
15 - Portalegre
16 - Faro
17 - C.Branco
18 - Bragança

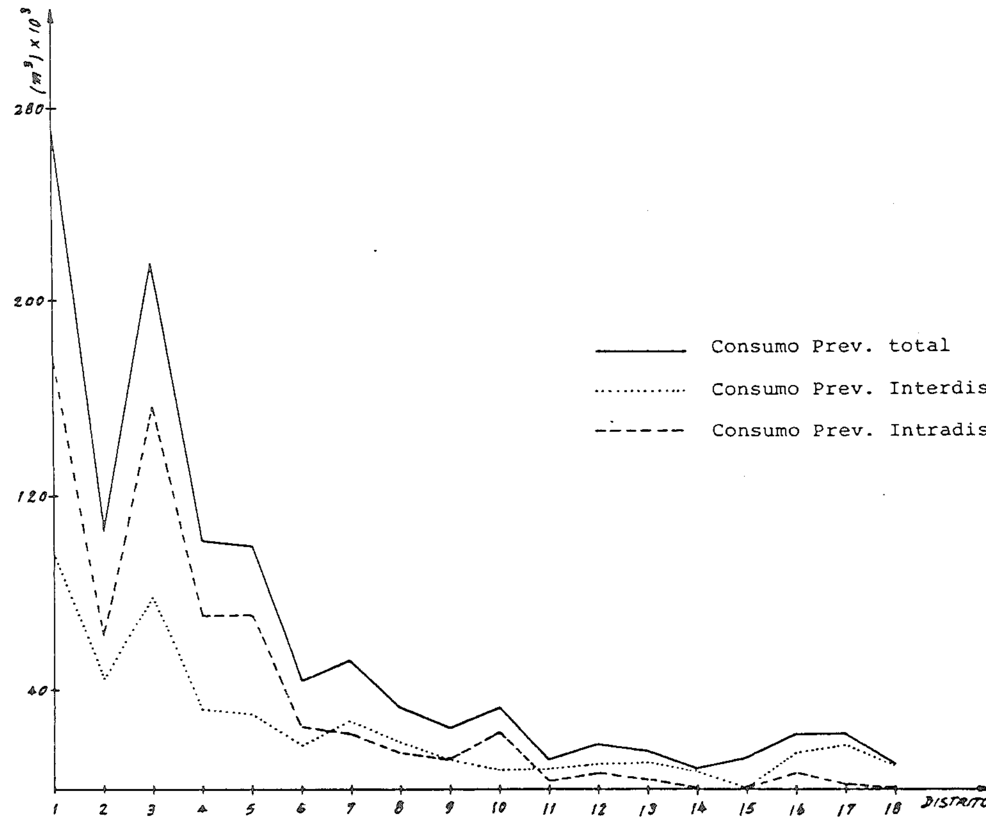


Fig. 5 - CONSUMO PREVISTO DE GASOLINA (1984) INTER E INTRADISTRITAL

QUADRO VII
CONSUMO DE GASOLINA NORMAL E SUPER PREVISTO
PARA 1984 (M^3) INTER E INTRA DISTRITAL

DISTRITO	CONSUMO TOTAL PREVISTO	CONSUMO INTER - DISTRITAL PREVISTO	CONSUMO INTRA - DISTRITAL PREVISTO
AVEIRO	102439	32044	70395
BEJA	8962	8679	283
BRAGA	101154	30119	71035
BRAGANÇA	10218	10112	106
C. BRANCO	20530	18096	2434
COIMBRA	43306	17736	25570
ÉVORA	13309	9582	3727
FARO	21640	15256	6384
GUARDA	15668	10931	4737
LEIRIA	51597	27826	23771
LISBOA	270479	92213	174266
PORTALEGRE	12501	11754	747
PORTO	237038	79794	157244
SANTARÉM	33663	18833	14830
SETÚBAL	106782	45413	61369
V. CASTELO	32028	8749	23279
V. REAL	17295	10319	6976
VISEU	24489	12625	11864
TOTAL	1123098	464081	659017

IV - Conclusões

Dos resultados atrás obtidos, poderemos concluir que:

a. O consumo total distrital de gasolina está relacionado com a riqueza gerada pela indústria e com o grau de ruralidade ou urbanidade do distrito.

b. O número de viagens interdistritais aumenta ou diminui dependendo das trocas comerciais efectuadas, podendo estas serem explicadas a partir da percentagem do VAB distrital para o total continental.

c. Mantem-se em 1984 a preponderância em consumo de gasolina dos distritos mais industrializados (Lisboa + Porto + Setúbal) (54.7%) acentuando-se a diferença que já era notória em 1980 (47.9%).

d. Os distritos mais pobres (Beja, Bragança e Portalegre) consumirão a maioria do combustível nas viagens interdistritais, facto este que, poderá encontrar explicação nas elevadas distâncias a percorrer até aos centros populacionais mais importantes, na generalidade não localizados no distrito.

e. A maior importância da parcela consumo intradistrital para aqueles distritos onde existe indústria implantada (Lisboa, Porto, Setúbal, Aveiro, Braga) traduzindo-se na utilização mais acentuada do automóvel privado em movimentos pendulares (por exemplo casa-trabalho).

f. Em relação ao total consumido as maiores percentagens do consumo de gasolina em viagens intradistritais situam-se nos distritos de Braga (70.2%), Aveiro (68.7%) seguindo-se Lisboa (64.4%) e o Porto (63.3%), facto que poderá ser explicado, dada a existência de infraestruturas no sector público dos transportes que permite uma utilização mais generalizada nos dois últimos.

Investigação Operacional, Vol.7 nº1 (Junho/1987) 121 - 126

O RISCO NA GESTÃO DE EMPRESAS: A FUNÇÃO DE UTILIDADE DE VON NEUMANN-MORGENSTERN

João Luís César das Neves
Luís Brito Frazão

Universidade Católica Portuguesa

Resumo: Pretende-se com este artigo ilustrar o interesse prático, na gestão corrente de uma empresa, do uso de uma função de utilidade face ao risco estimada para o gestor. Depois da apresentação dos pressupostos e resultados técnicos teóricos básicos de uma forma esquemática, são tecidas algumas considerações de carácter pragmático, terminando-se com a exposição do método de estimação ilustrada com os resultados obtidos para um gestor de uma empresa comercial.

Abstract: The purpose of this paper is to show the interest of the estimation of an risk-utility function for the manager of an enterprise. After presenting the hypothesis and theorectical results of the approach, the estimation of the function for a manager of a commercial firm is used to illustrate the procedure and the results.

Keywords: Risco, Teoria da Decisão, Função Utilidade, Estimação.

1 - A Tomada de Decisão em Universo de Risco

Quando postos perante a pergunta "Prefere receber 5.000$00 imediatamente ou que se deite uma moeda ao ar, se sair cara você ganha 10.000$00, se sair coroa não recebe nada?", a maioria das pessoas preferia a primeira hipótese! No entanto o valor esperado da segunda é exactamente o valor da primeira e portanto seria "lógico" que nos mostrassemos indiferentes perante elas.

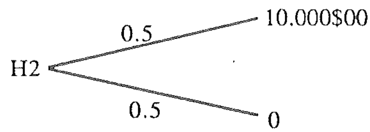
Esta "lógica" é uma falácia, visto que a segunda hipótese, de deitar a moeda ao ar, é substancialmente diferente da primeira: a possibilidade de receber 10.000$00 não é perfeitamente contrabalançada pelo risco de perder os 5.000$00, nada ganhando.

O que se passa é que, embora de valores esperados iguais, as duas hipóteses são diferentes num factor que sempre perturbou o homem, o risco!

O risco, presente no dia a dia da gestão de empresas, é algo de difícil abordagem, sujeito a idiossincracias e confusões de toda a espécie e portanto as decisões em situações de risco apresentam-se como um dos problemas mais "espinhosos" que a gestão de empresas contém.

Uma panóplia de métodos teóricos-empíricos, reunidos sob o título genérico de teoria da decisão, foi desenvolvida no sentido de estruturar e formalizar a tomada de decisão num universo onde a aleatoriedade domina, e, entre estes, a teoria da utilidade face ao risco tem um relevo justo devido ao carácter eminentemente subjectivo destas análises. Trata-se de esboçar, através de uma curva a atitude genérica de um indivíduo face a situações aleatórias.

A decisão descrita no início deste trabalho, é a escolha entre um valor certo (H1=5.000\$00) e uma "lotaria" (H2) ou seja um ganho dependente de uma variável aleatória.



Já vimos que não podemos substituir uma lotaria pelo seu valor esperado pois neste caso o valor esperado do jogo é igual a 5.000\$00 e só dificilmente poderiamos admitir a igual utilidade entre as duas hipóteses. Vimos, portanto, que a utilidade de H1(U(H1)) é maior que a utilidade esperada de H2:

$$U(H1) > U(E(H2)) \qquad (1)$$

ou seja, se substituirmos os ganhos pela sua utilidade, temos que:

$$E(U(10.000), U(0)) < U(E(10.000, 0)) \qquad (2)$$

representado por E(A, B) o valor esperado, calculado pelas probabilidades indicadas, entre A e B.

Mas qual é esta utilidade de que falamos e como a poderemos usar na análise ?

2 - A Função Utilidade de Von Neumann-Morgenstern

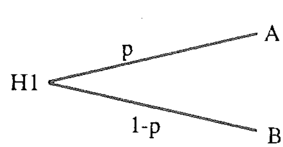
Uma das abordagens mais utilizadas para obter a utilidade dada por cada uma das situações em estudo é a construção de uma função que permita medir o valor de cada ganho numa escala ordinal de tal forma que se pudesse calcular a utilidade de uma lotaria entre A e B por

$$E(U(A), U(B)) \qquad (3)$$

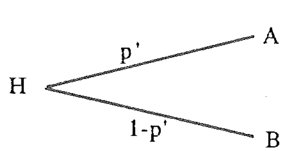
Para obter esta função é costume fazer um grupo de quatro hipóteses relativas às preferências, introduzidas na literatura por Von Neumann e Morgenstern que assim ficaram os patronos desta análise:

1º Axioma: É sempre possível estabelecer as preferências entre os prémios das lotarias, de uma forma clara. Não há dúvidas nem ambiguidades. Além disto as preferências são transitivas ou seja, se A é preferível a B, e B preferível a C, então A é preferível a C.

2º Axioma: É sempre possível definir preferências entre lotarias do mesmo modo que entre os seus prémios, ou seja, enfrentando duas lotarias:

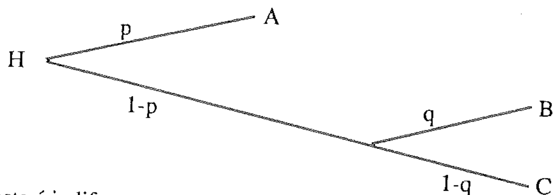


(4)



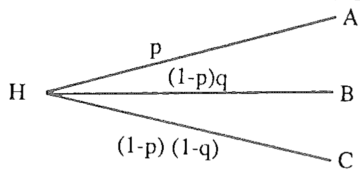
Se A é preferível a B, então H1 é preferível a H se e só se p maior que p'.

3º Axioma: Não há qualquer utilidade ganha pelo simples correr do jogo, ou seja, enfrentando uma lotaria que envolve dois jogos:



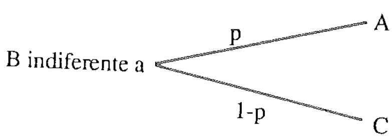
(5)

esta é indiferente a uma lotaria com apenas um jogo, com a apresentação seguinte



(6)

4º Axioma: As preferências são contínuas, o que é equivalente a supor que, dados três prémios quaisquer, A preferível a B e B preferível a C, existe uma probabilidade p e uma só, de tal forma que



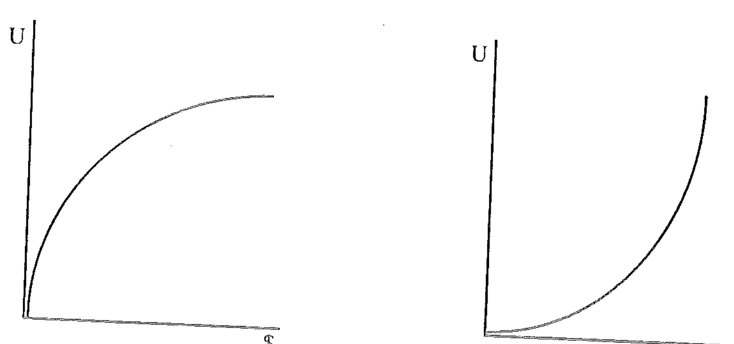
(7)

Não discutiremos nenhum destes axiomas que são sempre o ponto fraco de qualquer teoria.

Supondo um indivíduo ou um grupo com preferências que a eles obedeçam, é fácil demonstrar que estas são representáveis numa escala, a menos de uma transformação linear, sendo a função que relaciona esta escala com o valor monetário dos prémios a chamada "função utilidade do dinheiro face ao risco" de Von Neumann-Morgenstern, que garante que, supostos os axiomas:

$$U(A) > U(B) \iff A \text{ preferível a } B \qquad (8)$$

A atitude exposta no início deste artigo, da pessoa que posta perante uma lotaria com valor esperado igual a um ganho certo, prefere este último, é chamada de "aversão ao risco" dando à função de utilidade ùm aspecto semelhante ao da figura 1(a), enquanto a posição do "amante do risco" está esboçada em 1(b)



1(a) 1(b)

figura 1

3 - O Interesse da Estimação da Função de Utilidade

De posse dos cálculos que levam à estimação da função acima exposta para o gestor de uma empresa, a função de utilidade resume e explicita toda a atitude do indivíduo ou de um grupo face ao risco. Ela agora "toma decisões" como o faria o(s) entrevistado(s) perante uma ou várias lotarias.

Sendo dados os ganhos monetários e as probabilidades da situação em análise, a função aponta para a hipótese de maior utilidade para as preferências nela consubstanciadas.

Se se conseguir explicitar as preferências do conselho de administração de uma empresa, este pode delegar poderes num executivo que se obrigue a utilizar essa função em cada situação arriscada, ficando os membros do orgão dirigente certos de que as decisões serão as que eles tomariam.

Assim a delegação da tomada de decisão é um dos principais interesses desta análise.

Outra vantagem é a do analisado possuir um instrumento que retrata desapaixonadamente as suas atitudes perante o incerto. Quando confrontado com a situação real, a função pode servir de ponto de referência para a decisão, ponto de referência não influenciado pela tensão que a realidade pode trazer consigo.

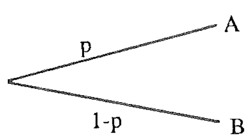
Assim é de todo o interesse que, numa empresa onde o risco represente uma componente importante da sua vida, o seu responsável procure objectivar o seu modo de tomada de decisão, com finalidade de poder delegar ou então com uso meramente indicativo.

4 - A Estimação da Função

A obtenção dos pontos pertencentes à curva torna-se simples com os axiomas acima expostos. Como foi dito a escala de utilidade é conseguida a menos de uma transformação linear, ou seja, detem dois graus de liberdade. Assim dois dos pontos podem ter valores de utilidade arbitrários, sendo esse o primeiro passo da estimação.

Tendo fixado a utilidade de A e B (por exemplo $U(A) = 0$ e $U(B) = 100$) os outros são conseguidos através de perguntas ao sujeito da experiência. Vários são os métodos da entrevista (ver referência 4) sendo os mais utilizados os de ir sucessivamente conseguinso obter a utilidade de novos valores à custa de lotarias com prémios de que já sabemos as utilidades.

Poder-se-á perguntar "qual a possibilidade p com que seria indiferente entre uma lotaria



(9)

e o ganho certo de C ?". Obtida a resposta, a utilidade de C seria conseguida à custa da relação

$$U(C) = p\,U(A) + (1 - p)\,U(B) \qquad (10)$$

Outra hipótese de fazer a mesma coisa seria definir p na lotaria (9) e perguntar o valor de C que seria indiferente a essa lotaria. Para ajudar são muitas vezes usados diagramas onde a percentagem p de certa figura (círculo, quadrado) está colorida, com o objectivo de "visualisar a probabilidade".

Muitos outros procedimentos são usados mas quase todos utilizam esta base, conhecida por "método do equivalente certo".

Vários cuidados têm de ser tomados na entrevista sendo os principais perigos os seguintes:

a) o efeito de aprendizagem por parte do entrevistado
b) o efeito de tomar muitas decisões
c) o efeito da ordem da colocação das perguntas
d) o refúgio do entrevistado em regras exógenas de decisão

Não discutiremos aqui estes riscos da análise, remetendo os interesses para o excelente trabalho empírico de Spetzler (ref. 4) que serviu aliás de base à estimação que se segue.

5 - Os Resultados

O estudo baseou-se em 22 questões do tipo acima exposto ou seja perguntando qual a probabilidade p de obter um lucro s (contra 1-p de obter um prejuízo r) para a qual esse negócio era julgado pelo entrevistado um negócio nem bom nem mau (indiferente a zero). Assim obtêm-se a utilidade de certos montantes que, adequando-se à situação da empresa em causa, representa um negócio com lucros que variavam entre 30 e 1000 contos e prejuízos de 25 a 600 contos. As várias situações foram postas em contraposição a zero ou seja não entrar no negócio.

Com estes valores procurou-se estimar a função utilidade que, à semelhança do estudo da referência 4, usou a especificação:

$$U(x) = A + B \log[\, x + C - D\,((x^2 + E^2)^{1/2} - E) \qquad (11)$$

onde $0 < D < 1$, $C > 0$ e $E > 0$.

Como a função tem de passar por dois pontos pré-fixados pode-se resolver a equação (11) em ordem a dois parâmetros (normalmente A e B) obtendo-se os outros através de um método de busca heurístico, de forma a minimizar o desvio quadrático:

$$\min \sum_i [U(0) - p(i)\, U(s(i)) + (1-p)\, U(r(i))]^2 \qquad (12)$$

onde o índice i indica as perguntas, feitas ao entrevistado, relativas a cada negócio.

Os cálculos foram efectuados num micro computador e a função obtida tem como parâmetros:

$A = -1633.1973$
$B = 181.5489$
$C = 8069.983$
$D = 0.9900$
$E = 195$

e resulta na figura 2.

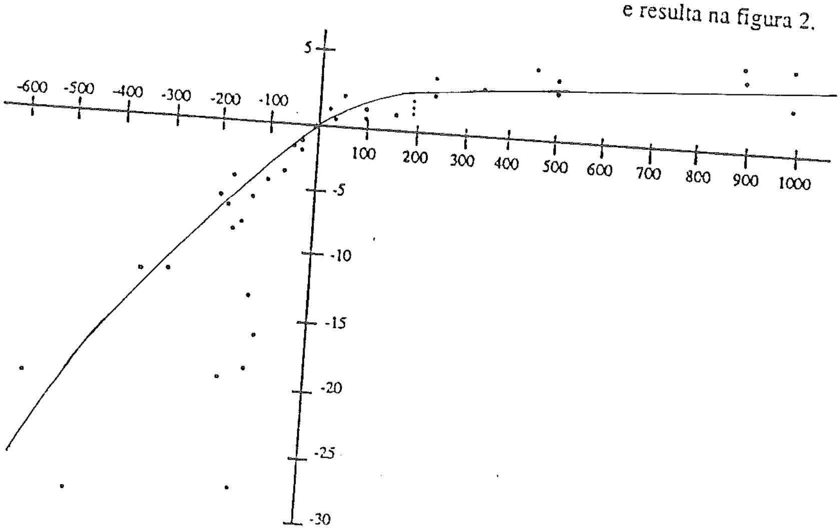


Figura 2

A forma de validar esta função proposta por Spelzler é a de calcular através do modelo os valores de p e compará-los com os p de indiferença dados pelo sujeito em estudo. Se os p estimados forem aceites pelo gestor então a função é boa.

Os valores de p calculados pelo modelo são os que impõem que

$$p\, U(s) + (1-p)\, U(r) = 0 \qquad (13)$$

ou seja,

$$p = -U(s)/[\, U(s) - U(r)\,] \qquad (14)$$

O quadro 1 dá alguns valores de p estimados pelo modelo e os obtidos pela entrevista, mostrando o grau de adequação do modelo à realidade.

Quadro 1

Negócio	p Entrevista	p Estimado	p Ent - p Est
1	0.77	0.7478	-0.0222
2	0.7	0.5191	0.1809
3	0.7	0.7750	-0.0750
4	0.48	0.4270	0.0530
5	0.95	0.8041	0.1459
6	0.48	0.6025	-0.1225
7	0.9	0.6351	0.2649
8	0.95	0.7487	0.2013
9	0.95	0.7574	0.1926
10	0.1	0.1773	-0.0773

6 - Conclusão

Procurou-se com este trabalho demonstrar o interesse e a simplicidade da análise da função utilidade face ao risco. De uma forma pouco dispendiosa e rápida poder-se-à obter um instrumento de gestão útil e prático, tornando menos arbitrária a vida da empresa.

Bibliografia

(1) Meyer, R.F. e Pratt, J. • "The Consistent Assessment and Fairing of Preference Functions", in IEEE Transactions on Systems Science and Cybernetics, Vol.SSC-4, Nº3, Sept.68, R. Howard(ed).

(2) North, D. W. • "A Tutorial Introduction to Decision Theory", in IEEE Transactions on Systems Science and Cybernetics, Vol.SSC-4, Nº3, Sept.68, R. Howard(ed).

(3) Raiffa, H. • "Decision Analysis: Introductory Lectures on Choices Under Uncertainty", Addison-Wesley, Massachusetts.

(4) Spelzler, C.S. • "The Development of a Corporate Risk Policy for Capital Investment Decisions", in IEEE Transactions on Systems Science and Cybernetics, Vol.SSC-4, Nº3, Sept.68, R. Howard(ed).

(5) Wilson, R.B. • "Decision Analysis in a Corporation", in IEEE Transactions on Systems Science and Cybernetics, Vol.SSC-4, Nº3, Sept.68, R. Howard(ed).

Fotografia, Montagem,
Impressão e Acabamentos
Secção de Textos da F.C.T.U.C.

ÍNDICE

Frederick J. Rigdway
Operations Research in Banking. Past Successes & Future Opportunities 3

Tom Bowen
Applications of OR in Public Services Planning 11

Heiner Muller-Merbach
Symmetric Codes for Revised Simplex Techniques 17

Francisco Amador Hidalgo, Manuel Cabanes Fuentes
La Aplication de la Investigation Operativa en Sistemas Educativos 27

Abílio Seca Teixeira
Modelo de Programação da Manutenção Preventiva de Grupos Térmicos Integrados num Sistema Electroprodutor Hídrico-Térmico 37

J. Rodrigues Dias
Um Algoritmo de Simulação-Optimização para Problemas Estocásticos 47

Fernando de Jesus
Uma Nota sobre a Função de Produção de Leontief 55

José Manuel Viegas
Redes de Serviços Públicos e Cobertura de Território 61

Jorge Guerreiro, Manuel Ramalhete
Dualidade e Contabilidade Analítica 67

J. Romão Eusébio, Lélio Amado
Modelos de Planeamento/Gestão para Optimização das Necessidades de Viaturas e Tripulações numa Rede de Transportes 81

Maria Natália Tavares, Victor Baptista
Descrição de um Modelo de Expansão do Sistema Energético 91

J. A. Assis Lopes
Metodologia na Modelação do Consumo Energético a Nível Distrital — Alguns Resultados . . . 109

João Luis César das Neves, Luis Brito Frazão
O Risco na Gestão de Empresas: A Função de Utilidade de Von Neumann-Morgenstern 121

apdio

Associação Portuguesa para o Desenvolvimento
da Investigação Operacional

CÉSUR — Instituto Superior Técnico — Avenida Rovisco Pais
1000 Lisboa — Telef. 80 74 55