# AN APPROACH TO DISTRIBUTED MODELING

Alexandr Umnov

June 1979
WP-79-48



INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

# PREFACE

The problem of computer linkage of different mathematical models into a whole system in order to investigate their joint behavior with common criterion and constraints seems to be taken more and more into consideration.

This paper presents a description of a particular approach which might be referred to as 'distributed modeling'. The main part of this paper was written by the author during his stay at IIASA in 1978.

SUMMARY

Many works on analyzing the behavior of complex systems are based on building large-scale integrated models and sequentials using decomposition and aggregation procedures. In this paper an approach is described which permits to investigate a set of linked subsystems without explicitly building any integrated model.

This approach based on the 'smooth' version of the Sequential Unconstrained Minimization Techniques (SUMT) can be considered from a mathematical point of view as a realization of the 'general decomposition scheme' (Orchard-Hays 1968).

Section 1 describes the statement of the problem; Section 2 gives a general description of the idea of the SUMT; Section 3 contains the conditions of applicability of the approach; and in Section 4 we give a short description of the practical realization and computer testing of this method for a case of linking two submodels of a health care system.

# AN APPROACH TO DISTRIBUTED MODELING

A. Umnov

## INTRODUCTION

Investigating an object of a complex structure, it is reasonable at the first stage of systems analysis to consider this object as a set of its independent parts. This way we can build mathematical models of all these parts at a sufficiently high level of detail. At the second stage, we have to take into consideration all interactions between the subsystem when they operate under common criteria and constraints.

All this raises the necessity of developing procedures, both methodological and computational, which give us the possibility of linking independent submodels into a whole system. There are two main approaches for solving this problem: the first is to design a large-scale integrated mathematical model describing the behavior of the system as a whole and consequent decomposition of the model. This method is very convenient from a computer point of view, as far as it needs comparatively low hard- and software resources. On the other hand, appropriate transformations of all files of the submodels into a common form are to be made in this approach. Besides, it is sometimes difficult to trace the process of optimization which can give information of a significant practical value.

The second approach consists of using one or several analysts or decision makers to organize interactions between submodels. This method permits us to link mathematical models without any file transformations, but it is practically impossible to use optimization procedures due to large time required per one iteration. Therefore it seems very desirable to develop an approach permitting linkage of different submodels in a direct way, without building a large-scale model which is to be partitioned later. We would also like this approach to enable linkage of submodels prepared independently by different groups of specialists and to give us the possibility of using different mathematical methods for solving the subproblems, perhaps on different computers. Briefly speaking, this approach must consider all submodels to be linked as 'black boxes' and assumes the possibility of using only their input and output data. The discussion of the approaches to the linkage of models and different economic applications are given in Bagrinovskii (1977). It should be noted that the second approach (which is generally more adequate to the 'linkage' problem) has been investigated in less detail than the first approach (usually associated with the decomposition of the problem.)

The purpose of this paper is to describe a scheme of realization of the second approach and to discuss some experience of its application.

## Statement of the Problem

The easiest way to link different submodels into a whole system consists of using special variables to formalize interrelations between submodels. These variables called further 'common variables' or 'coupling variables' will be denoted as V in contrast to 'inner variables' of the submodels which are denoted as X.

There are different ways of introducing these 'common variables' but we will now consider the general case omitting some details which will be discussed in other sections. Let us assume that each of the submodels can be formulated in terms of 'inner' variables as following

$$\text{minimize with respect to } X^k \qquad F^k(X^k) \quad ,$$

$$\text{subject to } X^k \in \mathfrak{E}^{n^k} \qquad G_s^k(X^k) \geq 0, \quad s = \overline{1,m}^k$$

$$k = \overline{1,N} \quad ,$$

where N is the number of submodels to be linked.

It is important to emphasize that all functions $F^k$ and $G_s^k$ are not known since we decided to consider the submodels as 'black boxes.'

After introducing 'common variables' in an appropriate way, we have the following statement for each subproblem

$$\begin{aligned} &\text{minimize with respect to } X^k \qquad F^k(X^k,V) \quad , \\ &\text{subject to } X^k \in \mathfrak{E}^{n^k} \qquad G_s^k(X^k,V) \geq 0, \; s = \overline{1,m}^k; \end{aligned} \tag{1}$$

where V is a vector of 'common variables', $V \in \mathfrak{E}^L$, and is fixed in (1).

The second step of linking these submodels consists in formalizing relations between them. Let these relations be given as a set of constraints on 'common variables'

$$R_s(V) \geq 0 \quad , \qquad s = \overline{1,M} \quad ,$$

where M is the number of these constraints.

Finally, we have to formulate the common criterion of operating the whole system of submodels to be linked. There are several reasons why we should use this criterion as a linear combination of criteria from different submodels, subject to all weight coefficients of this combination which are nonnegative. Some of the foundations of this choice will be discussed later on.

Therefore, we have the following system of relations to find optimal values of 'inner variables' $X^k$ and 'common variables' V:

$$\text{minimize with respect to } X^k \text{ and } V \quad \sum_{k=1}^{k=N} \lambda_k F^k(X^k,V)$$

$$\text{subject to} \quad G_s^k(X^k,V) \geq 0 \quad , \quad s = \overline{1,m}^k;\ k = \overline{1,N} \quad , \qquad (2)$$

$$R_s(V) \geq 0 \quad , \qquad s = \overline{1,M} \quad ,$$

where $\lambda_k$ are nonnegative weight coefficients.

Theoretically, a solution of this problem gives us all the desired data. But in the first place, it seems to be impossible to solve it, since $F^k$ and $G_s^k$ are not known to us and, second this problem is a very large one. In order to overcome these difficulties it is advisable to employ the software of the submodels which provides us with optimal values of 'inner variables' for fixed and perhaps nonoptimal values of 'common variables'.

Let $X^{*k}(V)$ be a solution of problem (1), subject to the vector of 'common variables' being fixed. Substituting this solution for each of the subproblems (1) to replace $X^k$ in (2), we get a new problem

$$\text{minimize with respect to } V \sum_{k=1}^{k=N} \lambda_k F^k(X^{*k}(V),V)$$

$$\text{subject to } V \in D \text{ and} \qquad (3)$$

$$R_s(V) \geq 0 \quad , \qquad s = \overline{1,M} \quad ,$$

where D is the domain of definition of functions $X^{*k}(V)$ .

Two explanations are necessary for the statement of problem (3) because it is the basic problem in our consideration. First, we have to take into account the domain of definition D, as the subproblems (1) do not have a feasible solution for any vector of 'common variables'. Secondly, 'inner constraints' $G_s^k \geq 0$ of (1) are omitted in (3) because they are satisfied by $X^{*k}(V)$ by definition. We shall further call problem (3) a 'master problem'.

By some natural assumptions on conditions of the subproblems we can find the optimal values of $X^k$ as $X^{*k}(V^*)$, where $V^*$ is the solution of the 'master problem' (3). This gives us the possibility of finding solutions $X^{*k}$ of the submodels independently after solving (3), which is a more preferable problem than (2) because of its less dimension. But from another point of view,

there are two difficulties peculiar to the 'master problem': to solve (3) we have to know the domain of the definition D and the explicit form of the functions $X^{*k}(V)$. It is unlikely that this sort of data will be found in most practical cases and we have to find an indirect way of solving the 'master problem'.

To surmount the difficulties mentioned above, we can take into account the fact that any numerical algorithm for solving a mathematical programming problem needs only some numerical data associated with the current approximation of the solution, but not the explicit form of the condition of the problem. In other words, for operating these algorithms we must be able to calculate only some numerical characteristics of functions $X^{*k}(V)$ such as their values and maybe their derivatives, at some points V.

As to the domain of definition D, we can avoid the necessity of explicitly building this set by using special procedures checking the existence of $X^{*k}(V)$ at any given point V, or by such kinds of algorithms which give the pseudo-solution of the problem when it has no feasible points. Therefore, the scheme of solving the whole problem can be formulated as following.

For some current point V in the space of 'common variables', we find all the data needed for solving the 'master problem'. It is likely that all the subproblems have to be solved for this fixed V. We then change the values of the 'common variables' according to the procedure of minimizing the common criterion of the whole problem. Repeating these two steps we eventually receive the optimal value of V. It is necessary to emphasize that this scheme is also considered as a variant of a 'general decomposition approach' described by W. Orchard-Hays (1968).

## General Description of the Approach

There are many works in which 'master problems' are used in different decompositional schemes (see, for example Fiacco and McCormick 1968 and Geoffrion 1970). In these works problems (1) and (3) are considered directly and that is the reason why we can't apply the standard algorithms of 'smooth' optimization to solve (3). The main difficulty preventing this is that

functions $X^{*k}(V)$ are not differentiable with respect to V. It makes us use special delicate methods of analyzing their properties if only problems (1) and (3) are considered in the form given in the previous section, and provides the idea of using nondifferentiable procedures of optimization to solve (3) (Lemarechal 1978).

But there is a way in which we can make computer linkage of different submodels on the basis of 'smooth' algorithms. The idea consists of a preliminary transformation of the problems (1) and (3) providing them with some desirable properties. This transformation is offered to be made according to the Sequential Unconstrained Minimization Techniques (SUMT), sometimes called the Penalty Functions Method as well.

This method (exactly speaking, its 'smooth exterior point' version (Fiacco and McCormick 1968) consists of unconstrained minimizations of some auxiliary function associated with the mathematical programming problem to be solved.

Let $E^k(X^k,V,T)$ be this auxiliary function for the kth problem (1), and $\hat{X}^k(V,T)$ be an extremal point of this function. Then, under some natural assumptions, the following relation between $\hat{X}^k(V,T)$ and $X^{*k}(V)$ takes place

$$\lim_{T \to +0} \hat{X}^k(V,T) = X^{*k}(V) \quad , \tag{4}$$

where T is a positive fixed parameter defining the degree of penalty for violations of constraints. This means that the extremal point of this auxiliary function is the solution of the problem (1) with perhaps some small error.

In the approach under consideration there are two reasons why it is convenient to use the 'exterior point' version of the SUMT. First, auxiliary functions for problem (1) will always have an extremal point independent on whether the problem has a feasible solution or not. Second, the 'smooth' version of the SUMT gives us the possibility to find all necessary data associated with $\hat{X}^k(V,T)$ by using a well known implicit function theorem (if, of course, all required derivatives exist.)

Let us choose the auxiliary function $E^k$ in the following form

$$E^k(X^k,V,T) = \lambda_K F^k(X^k,V) + \sum_{s=1}^{s=m^k} P(G_s^k(X^k,V),T) \qquad (5)$$

where the used penalty function $P(A,T)$ is defined and has continuous partial derivatives of the second order for any $T > 0$ and any A, and it satisfies the following relation as well

$$\lim_{T\to+0} P(A,T) = \begin{cases} 0, & \text{for any } A > 0 \\ +\infty, & \text{otherwise} \end{cases} \qquad (6)$$

An auxiliary function associated with the 'master problem' (3) can be chosen as

$$E(\hat{X}(V,T),V,T) = \sum_{k=1}^{k=N} \lambda_k F^k(\hat{X}^k(V,T),V) + \sum_{s=1}^{s=M} P(R_s(V),T) + \sum_{k=1}^{k=N} \sum_{s=1}^{s=m^k} P(G_s^k(\hat{X}^k(V,T),V),T) \quad . \qquad (7)$$

The double sum in this formula presents the penalty term for violations of the domain of definition D.

After obvious transformations we see that

$$E(\hat{X}(V,T),V,T) = W(V,T) + \sum_{k=1}^{k=N} E^k(\hat{X}^k(V,T),V,T) \qquad (8)$$

where

$$W(V,T) = \sum_{s=1}^{s=M} P(R_s(V),T) \quad .$$

Expression (8) is of great importance as it presents the auxiliary function (7) as a sum of the auxiliary functions associated with the problems (1) and a function W given in the explicit form.

Let $\hat{V}(T)$ be an extremal point of (7); then approximate values of the 'inner variables' can be given as $\hat{X}^k(\hat{V},T)$.

Our first problem is to give the procedure of finding $\hat{V}(T)$ and, second, to consider the problem of evaluating $V^*$ and $X^{*k}$. The problem of accuracy will be discussed later, and the main attention will now be paid to describing the data needed for solving the 'master problem' (3) by minimizing the auxiliary function (7).

By virtue of assumptions stated above, any standard scheme of unconstrained optimization may be used for finding V. As a rule, these schemes consist of building a sequence of points in the space of the 'common variables' $V^{(i)}$, which converges to V and is defined by the following recurrent equation

$$V^{(i+1)} = V^{(i)} + s\,Z^{(i)} \quad , \qquad i = 0,1,2,... \tag{9}$$

where $Z^{(i)}$ is a direction of minimizing (7), and s is an appropriate stepsize along this direction. Hence, there are two problems to be solved: how to find $Z^{(i)}$ and how to evaluate s.

In the first place, we shall consider the problem of building the direction of minimization for the auxiliary function (7). Doing this, we have to know the value, the gradient and perhaps the hessian matrix of the function to be minimized. We shall consider the case when all these data are needed in the chosen scheme of optimization.

Let GRAD and $\nabla_x$ denote conventional gradient operators in the spaces of 'common' and 'inner' variables respectively. In the same way, HESSIAN and $\nabla_x^2$ will be the hessian matrix with respect to 'common' and 'inner' variables.

It is very important to notice that all derivatives with respect to V have to take into account both explicit and implicit dependence of the function to be differentiated on 'common variables.'

By the chain rule

$$\text{GRAD } E = E'_V + \sum_{k=1}^{k=N} H^k_{xv}\,\nabla_x E^k \quad ,$$

where $E'_V$ is a vector of the partial derivatives of E with respect to V and $H^k_{xv}$ is the conventional matrix of sensitivity of $X^k$ by V.

Analogously,

$$\text{HESSIAN } E = E''_V + \sum_{k=1}^{k=N} H^k_{xv} E^{k}_{xv}{}'' + \sum_{k=1}^{k=N} H^k_{xvv} \nabla_x E^k$$

$$+ \sum_{k=1}^{k=N} H^k_{xv} (\nabla^2_x E^k (H^k_{xv})^t + E''^k_{xv}) \quad ,$$

where $E''_V$ is the matrix of second partial derivatives of E with respect to V, and

$E''^k_{xv}$ is $L \times n^k$ - dimensional matrix of partial derivatives of $E^k$ with respect to $X^k$ and V and, finally,
$H^k_{xvv}$ is the matrix of sensitivity of the second order.

Both these formulas are valid for any $X^k$ and V, but $\hat{X}^k(V,T)$ are the minimum points of the auxiliary functions $E^k$. Therefore, by virtue of the fact that $\hat{X}^k$ satisfies the following equation

$$\nabla_x E^k(\hat{X}^k, V, T) = 0 \quad , \qquad (10)$$

we simply have

$$\text{GRAD } E = E'_V \quad .$$

Taking into consideration (10) and that the full derivative of this equation with respect to V is

$$H^k_{xv} \nabla^2_x E^k + E''^k_{xv} = 0 \quad , \qquad (11)$$

we find

$$\text{HESSIAN } E = E''_V + \sum_{k=1}^{k=N} H^k_{xv} E''^k_{xv} \quad . \qquad (12)$$

To evaluate the quantity of information which is needed to transmit from each of the submodels to the 'master problem',

rewrite the expressions for GRAD and HESSIAN in new form

$$\text{GRAD } E = W'_v + \sum_{k=1}^{k=N} E^{k,}_v$$

and

$$\text{HESSIAN } E = W''_v + \sum_{k=1}^{k=N} (E''^{k}_v + H^k_{xv} E''^{k}_{xv}) \quad ,$$

by substituting (8).

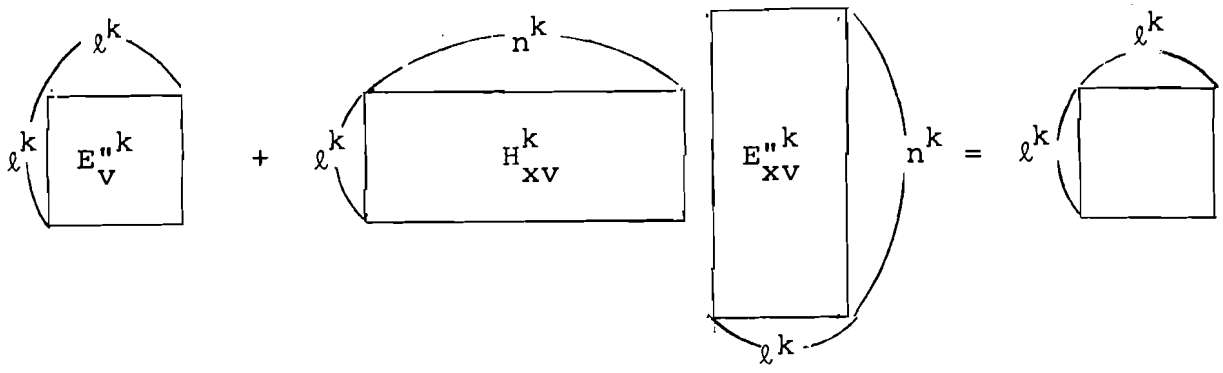
It is easy to verify that the matrices $E''^{k}_v + H^k_{xv} E''^{k}_{xv}$ are symmetrical. Actually, if we multiply both sides of (11) by $H^k_{xv}$ we get

$$H^k_{xv} E''^{k}_{xv} = - H^k_{xv} \nabla^2_x E^k (H^k_{xv})^t$$

and, by virtue of the symmetry of $\nabla^2_x E^k$ and $E''^{k}_v$, we find the desirable result.

Therefore, the vector $E^{k,}_v$ and the upper right half of the symmetrical matrix $E''^{k}_v + H^k_{xv} E''^{k}_{xv}$ are those which are to be calculated and transmitted by the submodels to the higher level of the whole system. Besides, there is no necessity of operating with any details of inner structure of the submodels to build the recurrent term of minimizing a sequence at the level of the 'master problem.'

Let the number of components for the vector of 'common variables' be used for joining the kth submodel and equal $l^k$. Then, considering that $n^k$ is the dimension of this very submodel, we can show the calculation procedure of the matrix $E''^{k}_v + H^k_{xv} E''^{k}_{xv}$ as following




It means that using the considered approach is only worthwhile if the dimensions of the subproblems are much greater than the numbers of components of V, belonging to the same subproblems, i.e.

$$n^k \gg \ell^k \quad .$$

It seems that this inequality may take place for many practical problems and, hence, the approach given above can be successfully used.

To complete the general description of this method it is necessary to note that the procedure of choosing the length of the step along the direction of minimization in the space of the 'common variables' s may be done according to any standard scheme of one-dimensional optimization or searching.

Some foundations of the method are given in the Appendix.

## Practical Realization and Computer Testing of the Approach

The approach described in the previous sections was implemented by the author on IIASA's PDP/11 to investigate the interactions of two small submodels of the Health Care Systems.

These submodels stated in the framework of DLP (Propoi 1976) present the development of manpower and technical capacity of the HCS. The first submodel describes the development of manpower and consists of two groups of state variables: 'number of specialists' and 'number of students'. The 'number of entered students' (enrollments) and the 'number of invited specialists'

(recruitments) are the control variables. The block-scheme for a fixed time period and one for the specialization of this submodel is shown in Figure 1.

The second submodel describing the development of technical capacity consists of three groups of 'state variables': 'hospital capacities', 'dispensary capacities' and 'drug production's capacities.' The 'increase of hospital capacities,' 'increase of dispensary capacities,' 'increase of drug production's capacities' and 'volume of drug import' are control variables. The block-scheme of this model is given in Figure 2.

The problem was to find a common resource allocation between these two submodels which is to be optimal in the sense of best satisfaction of demands of the population in the medical care. It means the criteria of operating these two submodels is to minimize the absolute values of differences between demands and current levels of supplying the population with all kinds of medical care.

Following the approach described in the previous sections we have to consider the common criterion of these submodels as a sum of their independent criteria, subject to the constraint which is the sum of common resources (limited and fixed for each of the time periods.) This means that we have a two-component vector of 'common variables', $V_1$ equals the volume of resources allocated for the 'manpower' submodel and $V_2$ equals the volume of resources allocated for the submodel of 'technical capacities'. Both state and control variables are considered as 'inner variables' for these submodels.

The scheme of linking the submodels under consideration into the whole system is shown in Figure 3.

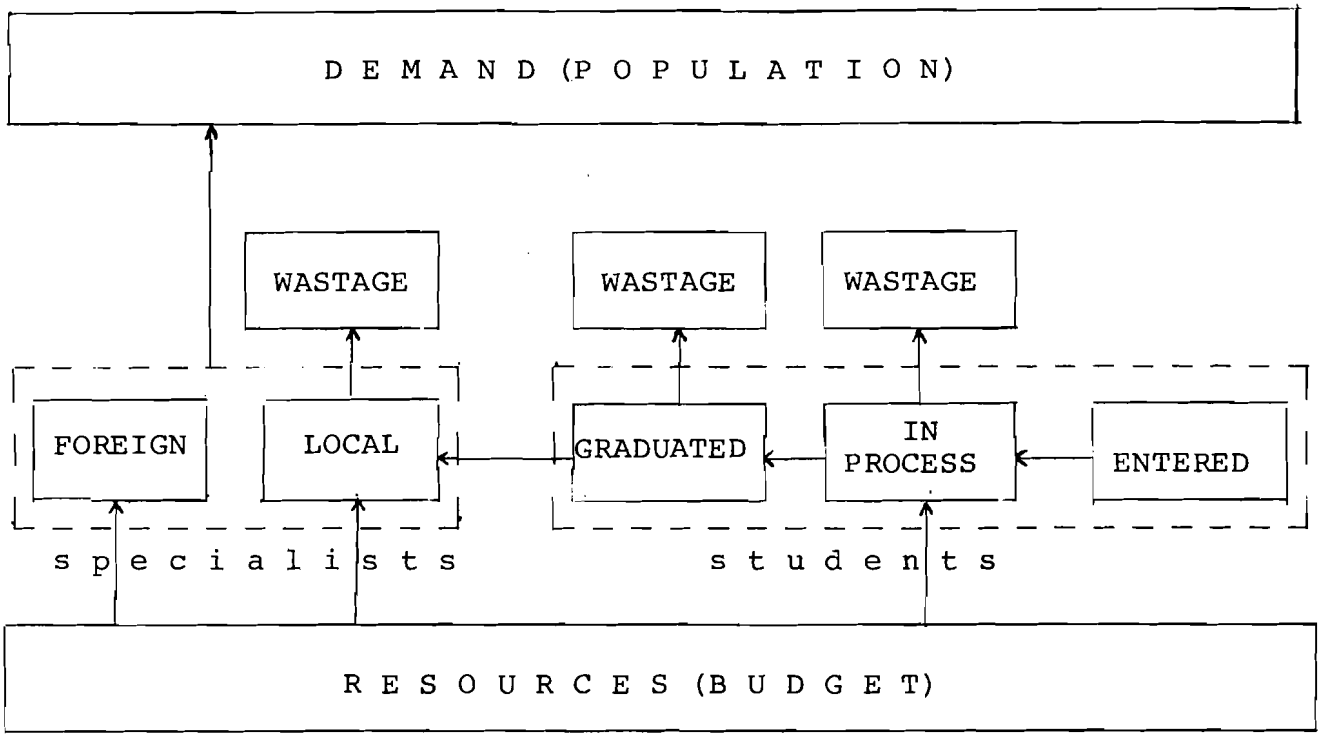



Figure 1.

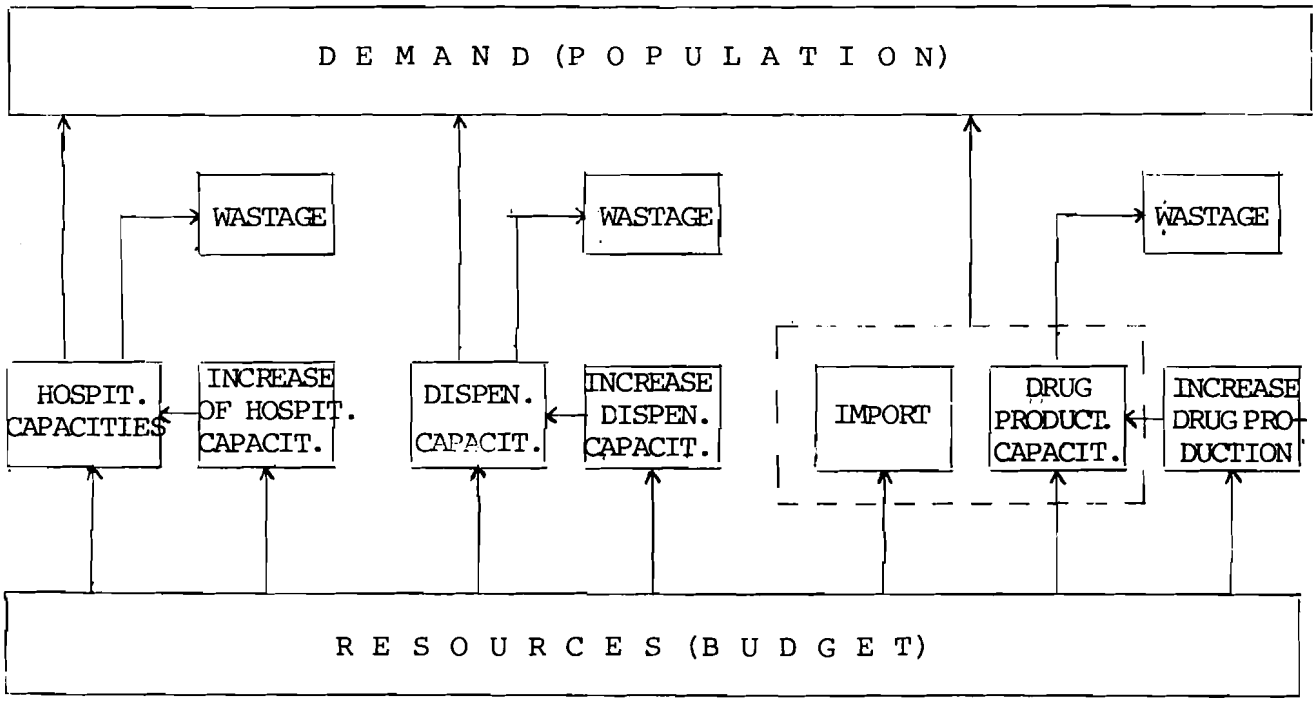



Figure 2.

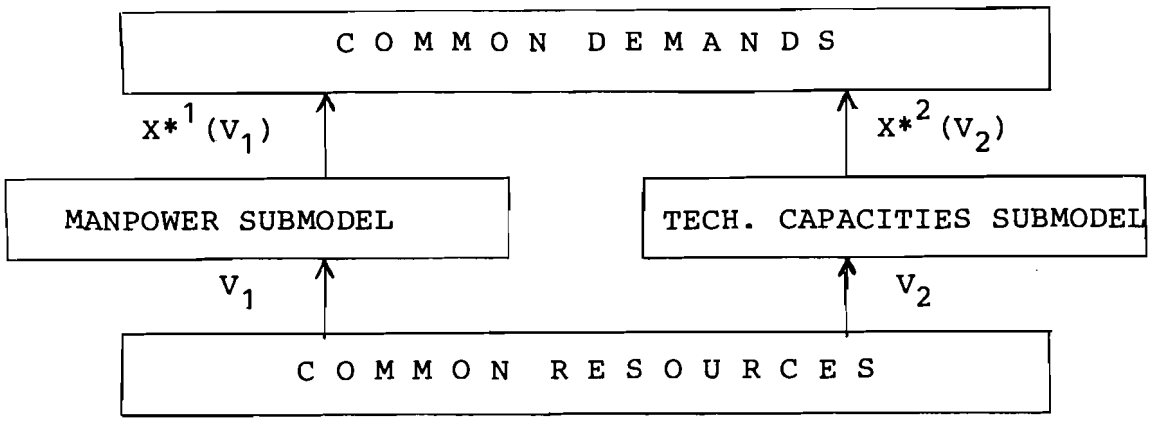



Figure 3.

The computer implementation of this problem was made on the 'smooth' version of the SUMT and the Newton-Raphson procedure. In this method the direction of minimization of the auxiliary function (7) is calculated by the following formula

$$Z^{(i)} = -(\text{HESSIAN } E^{(i)})^{-1} \text{GRAD } E^{(i)} \quad .$$

Considering the relatively small volume of the memory of the computer PDP/11 being available, the compact scheme of combination of the Newton-Raphson method and the conjugate gradient approach was used for finding $Z^{(i)}$. This scheme was developed and tested by A.G. Birukov at the Moscow Physical Technical Institute in 1974 (Birukov 1975).

Because it takes a lot of computational efforts to evaluate the optimal stepsize along the direction of minimization, the specific structure of the problem was taken into account to find s. It was the minimum of the length of the Newton vector and $\bar{s}$ was used as a stepsize, where $\bar{s}$ was the length of the step when only one nonactive constraint became active.

Two variants of the problem were solved for different initial data. One was the initial level of satisfaction of the population consisting of ten per cent of the demand and the other the level equalling to 90 percent. The optimal resources allocation is shown graphically in Figure 4.

DISTRIBUTION OF RESOURCES (in% )
IN SIX-PERIOD MODEL OF A HEALTH-
CARE SYSTEM

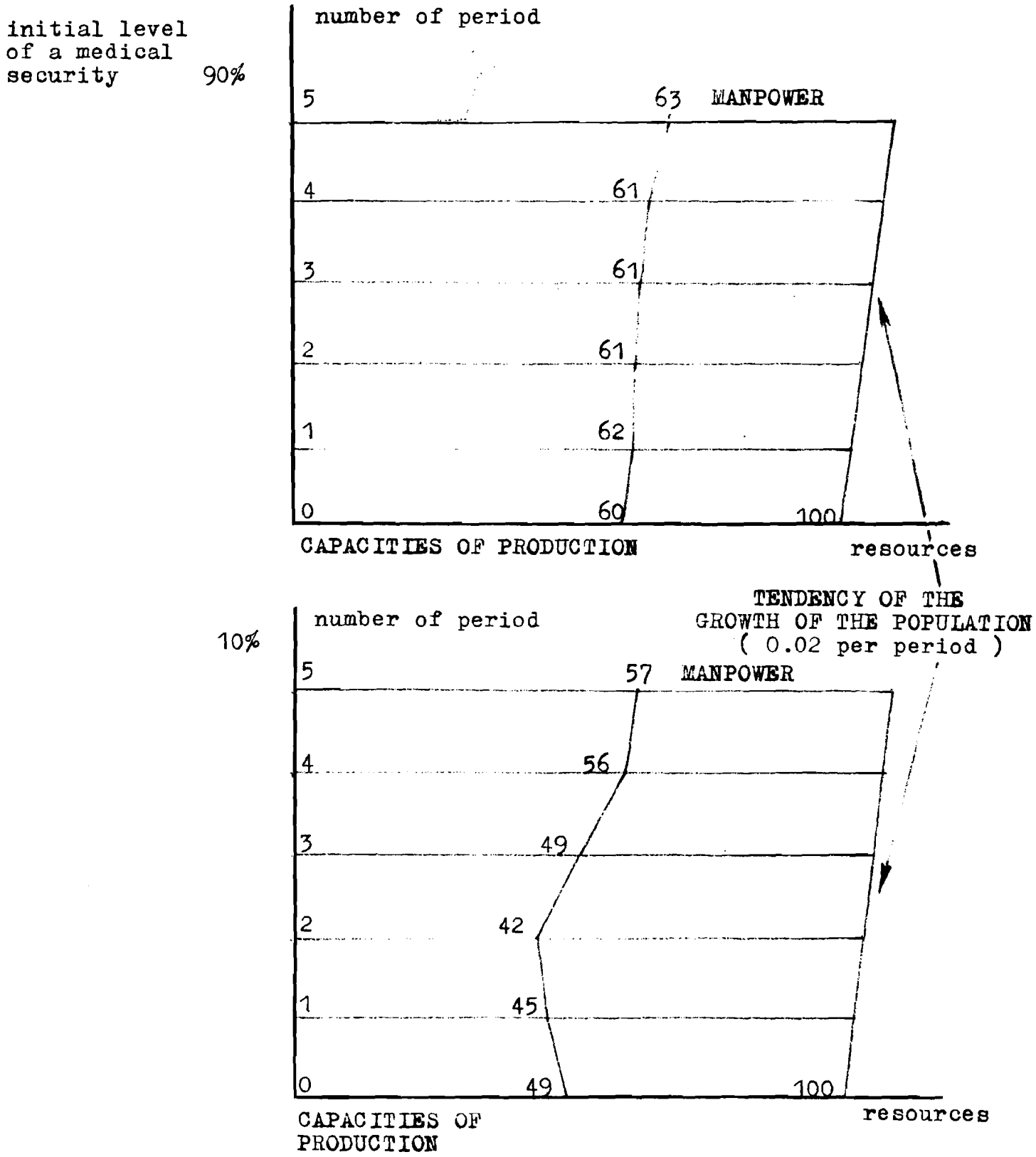



Figure 4

The difference in the results can be explained by the fact that the delay time in the manpower subsystem is twice as much than in another one. This means that we have to invest money in the first place in the manpower subsystem if only the initial level of the HCS is comparatively low.

## Conclusions and Some Further Developments of the Approach

The presented approach permits us to link different models into the whole system without explicitly building a large-scale mathematical programming model. This approach is opposed, in a methodological sense, to usual decomposition schemes, but is rather close to them from a mathematical viewpoint.

The main advantage of the approach is that all submodels are considered as 'black boxes' and therefore can be built by different groups of specialists on the base of different soft- and hardware.

The solution which may be found in this scheme is optimal in the sense of a new criterion which is a linear combination of the submodel's criteria with some nonnegative weight coefficients. It means that this approach can also be interpreted as a version of the multi-criteria optimization when a point of the Pareto set is a solution.

The approach allows different extensions and generalizations. We mention only two of them. It would be of interest to explore connections between linkage problems and multicriteria optimization because in both cases a man-machine procedure is appropriate. Second it is interesting to apply this approach to analysis of dynamic multistage optimization problems considering each stage as some local static optimization problems which are to be linked when the whole planning horizon is considered.

APPENDIX

In this appendix we shall study the conditions under which the considered scheme can be used.

Foundation of the Approach

Let functions $F^k$, $G_s^k$, and $R_s$ be 'smooth' enough and that problem (2) has an isolated local solution.

More accurately speaking, let functions $F^k$, $G_s^k$ and $R_s$ have continuous partial derivatives of the second order and let there be a system of points $V^*$ and $X^{*k}$ $(k = \overline{1,N})$ so that

$$G_s^k(X^{*k},V^*) \geq 0 \quad , \quad s = \overline{1,m^k} \quad ; \quad k = \overline{1,N}$$

$$R_s(V^*) \geq 0 \quad , \quad s = \overline{1,M} \quad , \tag{13}$$

there is a system of nonnegative numbers $p_s^k$ and $q_s$ satisfying the following relations

$$p_s^k G_s^k (X^{*k}, V^*) = 0 \quad \text{and} \quad p_s^k > 0$$

$$\text{if and only if } G_s^k(X^{*k},V^*) = 0 \quad , \tag{14}$$

$$q_s R_s(V^*) = 0 \quad \text{and} \quad q_s > 0$$

$$\text{if and only if } R_s(V^*) = 0 \quad \text{for all } s \text{ and } k.$$

Let U be the usual Lagrange function associated with the problem (2), i.e.

$$U = \sum_{k=1}^{k=N} \left(\lambda_k F^k(X^k,V) - \sum_{s=1}^{s=m^k} p_s^k G_s^k(X^k,V)\right) - \sum_{s=1}^{s=M} q_s R_s(V) \quad , \quad (15)$$

and at the point V*, $X^{*k}$ we have $\nabla_x U = 0$ and $\nabla_v U = 0$.

For any nonzero vectors $\Delta v$ and $\Delta x^k (k = \overline{1,N})$ such that

$$(\Delta v)^t \nabla_v G_s^k + \sum_{k=1}^{k=N} (\Delta x^k)^t \nabla_x G_s^k = 0, \text{ if } p_s^k > 0$$

and

$$(\Delta v)^t \nabla_v R_s = 0 \quad , \quad \text{if } q_s > 0 \quad \text{at the same point,}$$

the following unequality takes place

$$(\Delta v)^t \nabla_v^2 U \Delta v + 2 \sum_{k=1}^{k=N} (\Delta v)^t \nabla_{xv}^2 U \Delta x^k + \sum_{k=1}^{k=N} (\Delta x^k)^t \nabla_x^2 U \Delta x^k > 0 \quad (16)$$

where

$\nabla_v^2 U$ is the hessian matrix of U with respect to V,

$\nabla_x^2 U$ is the hessian matrix of U with respect to $X^k$,

and

$\nabla_{xv}^2 U$ is the matrix, elements of which equal to $\dfrac{\partial^2 U}{\partial x_k \partial v_i}$

then, by virtue of Theorem 4 [3] the point V* and $X^{*k}$ $(k = \overline{1,N})$ is an isolated local solution of the problem (2).

Under assumptions given above the following theorems will be valid.

*THEOREM 1.* $X^{*k}$ *is an isolated local solution of the problem (1) for fixed* V = V*.

*PROOF.* We have to show that all conditions analogous to (13 - 16) are valid for the problem (1) at the point $X^k$.

At first, we have by virtue of (13)

$$G_s^k(X^{*k},V^*) \geq 0 \quad , \qquad s = \overline{1,m}^k \quad .$$

It is possible to use the numbers $p_s^k$ defined in (14) as Lagrange-multipliers associated with the problem (1), then

$$p_s^k G_s^k(X^{*k},V^*) = 0$$

and

$$p_s^k > 0 \quad \text{if and only if} \quad G_s^k(X^{*k},V^*) = 0, \text{ for all } s,k.$$

By virtue of the separability of U with respect to $X^k$ we have

$$\nabla_x U^k = \nabla_x U = 0 \quad ,$$

where

$$U^k = \lambda_k F^k(X^k,V) - \sum_{s=1}^{s=m^k} p_s^k G_s^k(X^k,V) \quad .$$

Let $\Delta v$ and all $\Delta x^1,\Delta x^2,\ldots,\Delta x^N$ be equal to the zero vector except $\Delta x^k$, and for any nonzero $\Delta x^k$ from (16) we get that by $(\Delta x^k)^t \nabla_x G_s^k = 0$ when $p_s^k > 0$, the following unequality is valid

$$(\Delta x^k)^t \nabla_x^2 U^k \Delta x^k = (\Delta x^k)^t \nabla_x^2 U \Delta x^k > 0 \quad ,$$

and then all sufficient conditions of the optimality of $X^{*k}$ are proved. Q.E.D.

By virtue of the assumptions given above and a new assumption that all gradients of active constraints at $X^{*k}$ are linearly independent, we find from Theorem 6 [3] that functions $X^{*k}(V)$ exist within nonempty vicinity of V* and have at this point partial derivatives. It is necessary to notice that the existence of the matrix of sensitivity $H_{xv}^k$ does not ensure the differentiability of $X^{k*}(V)$ at V*.

Now we are able to prove

*THEOREM 2. The point V* is an isolated local solution of the problem (3).*

*PROOF*. By (13) and Theorem 1 we have

$$G_s^k(X^{*k}(V^*),V^*) = G_s^k(X^{*k},V^*) \geq 0 \quad ,$$

$$s = \overline{1,m^k} \quad ; \quad k = \overline{1,N}$$

and

$$R_s(V^*) \geq 0 \quad , \quad s = \overline{1,M} \quad .$$

Further, $p_s^k G_s^k(X^{*k}(V^*),V^*) = 0$, $p_s^k > 0$ if and only if $G_s^k(X^{*k}(V^*),V^*) = 0$ and $q_s R(V^*) = 0$, $q_s > 0$ if and only if $R_s(V^*) = 0$.

It means that nonnegative numbers $p_s^k$ and $q_s$ can be used as Lagrange-multipliers associated with the problem (3).

Let

$$\overline{U} = \sum_{k=1}^{k=N} (\lambda_k F^k(X^{*k}(V),V) - \sum_{s=1}^{s=m^k} p_s^k G_s^k(X^{*k}(V),V)) - \sum_{s=1}^{s=m} q_s R_s(V)$$

be the Lagrange function for problem (3). By the chain rule and Theorem 1 we have

$$\nabla_v \overline{U} = \nabla_v U + \sum_{k=1}^{k=N} H_{xv}^k \nabla_x U \quad ,$$

but granting (15), we find $\nabla_v U = 0$ at the point $V^*$.

Finally, let $\Delta v$ be any vector satisfying the following relations

$$(\Delta v)^t \nabla_v G_s^k(X^{*k}(V^*),V^*) = 0 \quad \text{if only} \quad p_s^k > 0$$

and

$$(\Delta v)^t \nabla_v R_s(V^*) = 0 \quad \text{if only } q_s > 0$$

at the point $V^*$.

Then, by virtue of the chain rule

$$(\Delta v)^t \nabla_v^2 \bar{U} \Delta v = (\Delta v)^t \nabla_v^2 U \Delta v + 2 \sum_{k=1}^{k=N} (\Delta v)^t \nabla_{xv}^2 U H_{xv}^k \Delta v$$

$$+ \sum_{k=1}^{k=N} (H_{xv}^k)^t (\Delta v)^t \nabla_x^2 U H_{xv}^k \Delta v \quad .$$

Denoting $\Delta x^k = H_{xv}^k \Delta v$ we get

$$(\Delta v)^t \nabla_v^2 \bar{U} \Delta v = (\Delta v)^t \nabla_v^2 U \Delta v + 2 \sum_{k=1}^{k=N} (\Delta v)^t \nabla_{xv}^2 U \Delta x^k$$

$$+ \sum_{k=1}^{k=N} (\Delta x^k)^t \nabla_x^2 U \Delta x^k \quad .$$

From another hand, new vector of local variations $\Delta x^k$ satisfies the following relations

$$(\Delta x^k)^t \nabla_x G_s^k(X^k, V) = 0 \quad \text{if only} \quad p_s^k > 0 \quad .$$

Really,

$$(\Delta x^k)^t \nabla_x G_s^k = (\Delta v)^t (H_{xv}^k)^t \nabla_x G_s^k = (\Delta v)^t \nabla_v G_s^k = 0 \quad ,$$

by virtue of our assumption.

Hence, $(\Delta v)^t \nabla_v^2 \bar{U} \Delta v > 0$ and V* is an isolated local solution of the 'master problem' (3). Q.E.D.

To finish the foundation of the approach we have to consider the properties of the penalty function P(A,T).

Except for the conditions stated above, this function will satisfy the following relations at any point from its domain of definition

$$\frac{\partial P}{\partial A} < 0 \quad \text{and} \quad \frac{\partial^2 P}{\partial A^2} > 0 \quad .$$

Then Theorem 10 [3] takes place and we have

$$\lim_{T \to +0} \hat{V}(T) = V^* \quad . \tag{17}$$

The possibility of using some of classical optimizational procedures for minimizing (7) arises from Theorem 2 and the assumptions about the existence of continuous partial derivatives of the second order for $F^k$, $G_s^k$, and $R_s$.

The Problem of Accuracy

Since the 'smooth' version of the 'exterior point unconstrained minimization techniques' gives us only the approximate solution of the problem to be solved, we have to consider the problem or find the exact solution or, at least, reduce the error of the approximation.

There are two aspects of the problem of accuracy in the given approach. First, we have no troubles because the solutions of (1) are approximate during the minimizational process of the auxiliary function (7), as this process is an iterative one. Second, we have to study the problem of approximation at the final point of the algorithm used.

The simplest way to evaluate the exact solution of the 'master problem' is to use the relation (17) and standard Taylor approximation of the function V(T).

Granting that

$$\hat{V}(T + \Delta T) = \hat{V}(T) + \Delta T \, \hat{V}'_T + o(T) \quad ,$$

where

$$\lim_{\Delta T \to 0} \frac{o(\Delta T)}{\Delta T} = 0 \quad .$$

Going over to the limit when $\Delta T \to -T$, we find

$$V^* = \hat{V}(T) - T \, \hat{V}'_T + o(T) \tag{18}$$

It means that to eliminate the linear part of the error we have to find the derivative $\hat{V}'_T$.

Before going into detail of this procedure it is necessary to notice that the assumptions being made in the previous section guarantee the existence of the trajectory of local minima of the SUMT and the validation of (18). Moreover, in [9] it is shown that if we choose the penalty function P(A,T) as a function of the single argument A/T then $\hat{V}'_T$ will have a bounded limit value by $T \to +0$.

Taking into consideration that the function $\hat{V}(T)$ is implicitly defined by the equation

$$\nabla_V E(\hat{X}(\hat{V},T),\hat{V},T) = 0 \qquad (19)$$

and by virtue of the implicit function theorem we get

$$\hat{V}'_T = (\text{HESSIAN } E)^{-1} \&''_{VT} \quad ,$$

where HESSIAN E is given by (12) and $\&''_{VT}$ is the derivative of the left part of (19) with respect to T.

As far as there are dependencies of $\nabla_V E$ on T both in explicit and implicit ways we have

$$\&''_{VT} = E''_{VT} + \sum_{k=1}^{k=N} \frac{\partial \hat{X}^k}{\partial T} \, t\, E''^{k}_{XV} + \sum_{k=1}^{k=N} H^k_{XVV} \nabla_X E^k$$

$$+ \sum_{k=1}^{k=N} H^k_{XV} \left(\nabla^2_X E^k \frac{\partial X^k}{\partial T} + E''^{k}_{XT}\right)$$

where

$E''_{VT}$ is a vector of partial derivatives of E with respect to V and T,

$E''^{k}_{XT}$ are vectors of partial derivatives of $E^k$ with respect to $X^k$ and T.

This formula can be rewritten as

$$\&''_{vT} = E''_{vT} + \sum_{k=1}^{k=N} H^k_{xv} E''^k_{xT} + \sum_{k=1}^{k=N} H^k_{xvv} \nabla_x E^k$$

$$+ \sum_{k=1}^{k=N} \frac{\partial \hat{X}^k}{\partial T}^t (\nabla_x^2 E^k (H^k_{xv})^t + E''^k_{xv}) \quad ,$$

but, by virtue of (10) and (11), we finally can find

$$\&''_{vT} = E''_{vT} + \sum_{k=1}^{k=N} H^k_{xv} E''^k_{xv} \quad .$$

It is followed by the fact that $\&''_{vT}$ can be calculated separately by different submodels. In fact, we have

$$\&''_{vT} = W''_{vT} + \sum_{k=1}^{k=N} (E''^k_{vT} + H^k_{xv} E''^k_{xT}) \quad .$$

Now we have to consider the problem of finding $X^{*k}$. Since this point is a limit one for $\hat{X}^k(\hat{V},T)$ when $T \to +0$, we can use the Taylor approximation again.

$$\hat{X}^k(\hat{V}+\Delta V, T+\Delta T) = \hat{X}^k(\hat{V},T) + \frac{\partial X^k}{\partial T} \Delta T + (H^k_{xV})^t \Delta V + o(\Delta V, \Delta T) \quad ,$$

where $\frac{\partial \hat{X}^k}{\partial T}$ can be found by means of the implicit function theorem from the equation (10).

Taking into consideration that $\Delta V = \hat{V}'_T \Delta T$ and going over to the limit when $\Delta T \to -T$, we get

$$X^{*k} = \hat{X}^k(\hat{V},T) - T \frac{\partial \hat{X}^k}{\partial T} - T H^k_{xV} \hat{V}'_T + o(T) \quad . \qquad (20)$$

This also means that the correction of the approximate solutions can be made independently by different submodels, but only after finding $\hat{V}'_T$ in the 'master problem.'

It can happen sometimes that one step of the procedure (18-20) doesn't provide us with the desirable level of accuracy. It is possible in this case to repeat all these calculations. The conditions of the convergence of the process which can be called 'iterative linear extrapolation' are given in [9].

## References


[1] Bagrinovskii, K.A, *Foundations of Planning Decision Coordinations*, Nauka, Moscow, 1977.

[2] Birukov, A.G., "A Decompositional Scheme of Solving an Unconstrained Minimization Problem," *Trudy MFTI*, (in Russian), 220-227, 1975.

[3] Fiacco, A.V., and McCormick, G.P., *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, J. Wiley, New York, 1968.

[4] Geoffrion, A.M., "Primal Resource - Directive Approaches for Optimizing Nonlinear Decomposable Systems, *Operations Research*, 18, 3(1970), 375-403.

[5] Lasdon, L.S., *Optimization Methods for Large Scale Problems*, MacMillan, London, 1970.

[6] Lemarechal, C., *Nonsmooth Optimization and Descent Methods*, RR-78-4, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1978.

[7] Orchard-Hays, W., *Advanced Linear Programming Computing Techniques*, McGraw-Hill, New York, 1968.

[8] Propoi, A.I., *Problems of Dynamic Linear Programming*, RM-76-78, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1976.

[9] Umnov, A.E., "The Iterative Linear Extrapolation in the Penalty Functions Method," *Jour.Comp.Math.&Math.Phys.*, 6, (in Russian), 1974.