NONDIFFERENTIABLE OPTIMIZATION PROMOTES HEALTH CARE

David Hughes
Evgeni Nurminski
Geoff Royston

October 1979
WP-79-90

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

FOREWORD

The principal aim of health care research at IIASA has been to develop a family of submodels of national health care systems for use by health service planners. The modeling work is proceeding along the lines proposed in the Institute's current Research Plan. It involves the construction of linked submodels dealing with population, disease prevalence, resource need, resource allocation, and resource supply.

This paper is an output of a collaboration between two Areas at IIASA. It describes how a health resource allocation model, developed in the Health Care Systems Task of the Human Settlements and Services Area, may be solved by using non-differentiable optimization techniques studied in the Optimization Task of the System and Decision Sciences Area.

Related publications in Health Care Systems and in Non-differentiable Optimization are listed at the end of this report.

Andrei Rogers
Chairman
Human Settlements
and Services Area

Andrzej Wierzbicki
Chairman
System and Decision
Sciences Area

ABSTRACT

An example of a health resource allocation model, solved previously by piecewise linear approximation with data from Devon, U.K., is solved using nondifferentiable optimization (NDO). The example illustrates a new application for NDO, and the novel approach makes clearer the workings of the model.

CONTENTS

NONDIFFERENTIABLE OPTIMIZATION PROMOTES HEALTH CARE

David Hughes, Evgeni Nurminski, and Geoff Royston

1. INTRODUCTION

Health care systems (HCS) and nondifferentiable optimization (NDO) are both studied at IIASA. Those who study HCS (like the first author) seek to model the features of health care systems that are common to different countries, so as to assist those who plan health services. The mathematicians interested in NDO (like the second author) seek to extend the classical optimization techniques to functions that have "nonsmooth" regions where no unique gradient can be defined. Shigan et al (1979) describe recent progress at IIASA in HCS modelling. The papers from a recent IIASA workshop on NDO were brought together by Lemarechal and Mifflin (1978).

This paper reports how a health resource allocation model used by the third author was solved by minimizing a function with points of nondifferentiability. Section 2 describes how an example of the model arose in the joint strategic planning of health and personal social services in Devon, U.K., a county with a population of about 1 million. Section 3 formulates the model as a problem for NDO. Ways to obtain numerical solutions are reviewed in Section 4 which compares the solution of the example by NDO and by another method based on linear approximation. Section 5 concludes.

2. RESOURCE ALLOCATION MODELLING IN DEVON

Devon is an area in the southwest of England, in which health services (e.g. hospitals, clinics) are managed by the Area Health Authority (AHA), and personal social services (e.g. residential homes, social workers) are managed by the Local Authority (LA). Many individuals receive both sorts of services which often overlap. After surgery, for example, some hospital patients may be discharged earlier if suitable nursing support is available for them at home. Elderly people may receive equivalent care in residential homes or in geriatric hospitals. The problem for Devon is to provide a balanced mix of health and personal social services within constraints on total resources.

McDonald et al (1974) describe a model to help in this task. It models the balance chosen by the many agents in the HCS (doctors, nurses, social workers, etc.) between the use of health services and personal social services for different categories of patients. The model's underlying hypothesis is that the aggregate behavior of these agents can be represented as the maximization of a utility (or inferred worth) function, whose parameters can be estimated from the results of previous choices. If these parameters do not change with time, the model can be used to simulate how future resource levels will be allocated in the HCS. Furthermore, because the underlying hypothesis is an optimistic one, the model may suggest reallocations. The full model is quite sophisticated with several special features. Only a simple version is reported here, both to clarify the presentation and because the example is one that actually arose in using the model to assist health care planning in Devon.

Table 1 categorizes elderly patients (65 or older) under 17 headings according to their housing, social isolation, physical disability, mobility and mental state. This categorization is part of a more detailed classification designed in conjunction with case workers who meet the patients. Table 2 lists 6 resources used in the domiciliary care of these patients. The first two resources (psychiatric and geriatric day hospitals)

Table 1. Seventeen categories of elderly patients.

	Defining Factors				
Patient category	Housing condition (1)	Social isolation (2)	Physical disability (3)	Mobility (4)	Degree of dementia in mental state
1	poor/good	mild	very severe	severe/mild	severe
2	poor/good	mild	severe	mild	severe
3	poor/good	mild	very severe	severe/mild	mild
4	poor/good	severe	severe	mild/good	mild
5	poor/good	mild	severe	mild/good	mild
6	poor/good	severe	mild	mild/good	mild
7	poor/good	mild	mild	mild/good	mild
8	poor	mild	very severe	severe/mild	none
9	good	mild	very severe	severe/mild	none
10	poor	severe	severe	mild/good	none
11	good	severe	severe	mild/good	none
12	poor	mild	severe	mild/good	none
13	good	mild	severe	mild/good	none
14	poor	severe	mild	mild/good	none
15	good	severe	mild	mild/good	none
16	poor	mild	mild	mild/good	none
17	good	mild	mild	mild/good	none

(1) Good housing means easy access to inside toilet and hot water. Poor housing means neither.
(2) Mild social isolation means not living alone. Severe social isolation means living alone.
(3) Mild - unable to carry out household care. Severe - unable to carry out household and personal care. Very severe - incontinent and/or unable to feed.
(4) Mild - can get around house, or can get out of house with aids or personal assistance. Severe - chairfast or bedfast.

Table 2. Six resources for domiciliary care.

Name of resource	Unit of resource
Psychiatric day hospital	day place
Geriatric day hospital	day place
Home nurse	WTE[1]
Day center	place
Home help	WTE[1]
Meals	service

[1]WTE = whole time equivalent (many nurses work only part-time).

are provided by the AHA; the others by the LA. Other institutional resources (such as in-patient hospitals and residential homes) are also used by elderly patients in Devon, but for this exercise their use was supposed to be fixed.

Patients in each of the 17 categories could receive many different combinations of the 6 resources. Table 3, however, defines up to 4 alternative modes of care for each category. These alternatives, which derive from discussions with consultants, senior nurses and other professionals, indicate how much of each resource might be used to provide equivalent levels of care for each patient. In a sense, the resource levels in these alternative "packages" represent ideal standards which doctors would like to attain. Unfortunately, these standards lie well above what can currently be afforded. Devon AHA and Devon LA want together to provide a mix of health and personal social services which they can afford and with which the HCS can approach the ideal standards for a large number of patients. The model was used to assist this debate by simulating who gets what.

In order to set up some mathematics, we use the indices

$i = 1,2 \ldots 17$ patient categories

$k = 1,2 \ldots 6$ resource types

$\ell = 1,2 \ldots 4$ care modes

and label the numbers in Table 3 as

$U_{ik\ell}$ = the ideal levels of resource type k in care mode ℓ for patient category i.

Because of resource constraints, rather lower resource levels $u_{ik\ell}$ are actually achieved, and it is these that the model seeks to predict. It also predicts

$x_{i\ell}$ = the numbers of patients in category i who receive care in mode ℓ

Table 3. Resources needed by elderly patients in alternative modes of care.

Patient category as defined in Table 1	Mode of care	Amount of resource needed per patient per year[1]					
		Psych. day hospitals	Geriatric day hospitals	Home nurse	Day center	Home help	Meals
1	1			1125	85	235	120
	2		100	1125		220	105
	3	100		1125		220	105
2	1	200		330		65	
	2			540	105	110	
	3			770		155	100
3	1		150	520		175	65
	2			690	85	235	120
	3			910		310	205
4	1		125	165		530	25
	2	250		75		235	
	3			165	125	530	25
	4			255		825	150
5	1		100	105		100	
	2			105	105	100	
	3			150		145	100
6	1	250				80	
	2				50	245	50
	3					285	100
7	1	50				40	10
	2		50			40	10
	3				40	40	10
	4					45	50
8	1		150	490		200	55
	2			860		330	205
9	1		150	490		150	55
	2			860		270	205
10	1		125	100		560	25
	2			100	125	560	25
	3			155		870	150

continued

Table 3 continued

Patient category as defined in Table 1	Mode of care	Amount of resource needed per patient per year[1]					
		Psych. day hospitals	Geriatric day hospitals	Home nurse	Day center	Home help	Meals
11	1		100	110		555	50
	2			100	125	500	25
	3			155		810	150
12	1		75	75		145	25
	2			65	105	130	
	3			90		185	100
13	1		50	80		85	50
	2			65	105	70	
	3			90		125	100
14	1		50			275	50
	2				50	275	50
	3					320	100
15	1		50			215	50
	2				50	215	50
	3					260	100
16	1		50			70	
	2					80	50
17	1		50			70	
	2				40	70	10
	3					20	50

[1]The units in this table are (for each resource respectively):
- daily attendances (1 psychiatric day place = 500 daily attendances.
- daily attendances (1 geriatric day place = 1000 daily attendances.
- visits (1 home nurse WTE = 3820 visits).
- daily attendances (1 day center place = 125 daily attendances).
- hours (1 home help/WTE = 1550 hours
- meals (1 meals service = 1000 meals).

so as to satisfy constraints on the total numbers d_i of patients in each category receiving care, and the total resources A_k of each type available for care,

$$\sum_{\ell} x_{i\ell} = d_i, \qquad \forall\ i \tag{1}$$

$$\sum_{i,\ell} x_{i\ell} u_{ik\ell} = A_k, \qquad \forall\ k \tag{2}$$

Both d_i and A_k are assumed to be known, and Tables 4 and 5 give the numbers of elderly patients, and the levels of health service and personal social service resources, used in the Devon example. The former arise from assuming that an approximately constant proportion of the elderly need care; the latter from certain assumptions about growth in the U.K. health service.

It remains to specify the form of the utility function maximized by the model. It is

$$Z(\underline{x},\underline{u}) = \sum_{i,k,\ell} x_{i\ell} h_{ik\ell}(u_{ik\ell}) \tag{3}$$

where

$$h_{ik\ell}(u) = \frac{C_k U_{ik\ell}}{\beta_k} \left[\left(\frac{u}{U_{ik\ell}} \right)^{\beta_k} - 1 \right] \tag{4}$$

$$\beta_k = 1 - 1/F_k \tag{5}$$

and where $\underline{x},\underline{u}$ denote $\{x_{i\ell},\ i = 1,2,\ldots,17,\ \ell = 1,2,\ldots,4\}$, $\{u_{ik\ell},\ i = 1,2,\ldots,17,\ k = 1,2,\ldots,6,\ \ell = 1,2,\ldots,4\}$, respectively. The funciton $Z(\underline{x},\underline{u})$ is

1) additive across i,k,ℓ. This implies no correlation between the objectives of increasing each and every $x_{i\ell} h_{ik\ell}(u_{ik\ell})$.

2) linearly increasing in $x_{i\ell}$. The extra benefit from taking care of one more patient in a particular care mode is independent of the number already cared for in that mode.

Table 4. Number of elderly in Devon.

Patient category i	Number of elderly patients d_i	Patient category i	Number of elderly patients d_i
1	43	10	51
2	38	11	198
3	326	12	132
4	90	13	777
5	200	14	918
6	891	15	3410
7	703	16	339
8	184	17	2667
9	818		

Table 5. Model parameters for example.

Resource type k	Resources available in Devon (units as in Table 2) A_k	Resource costs (£ running per year) C_k	Elasticities (as defined in equation (5)) F_k
1	10	5830	0.595
2	30	9190	0.800
3	125	5665	0.800
4	79.6	374	0.202
5	773.5	1778	0.325
6	275.4	230	0.325

3) zero when $u_{ik\ell}$ equals $U_{ik\ell}$ for all i,k,ℓ. At this point, marginal increases in Z resulting from increasing resource levels equal the marginal resource costs C_k. Normally, $u_{ik\ell} < U_{ik\ell}$ for some i,k,ℓ, and Z is then negative.

4) monotonically increasing and concave downwards in $u_{ik\ell}$ for $\beta_k \leq 0$. This implies diminishing returns as the ideal resource standards are approached. The speed with which the returns diminish is measured by the power parameters β_k, or the corresponding elasticities F_k.

5) not unlike a similar function defined in the model DRAM (Hughes and Wierzbicki, 1979). DRAM, however, does not incorporate the constraint (1) and does not require NDO.

Whether the results of maximizing the function $Z(\underline{x},\underline{u})$, subject to the constraints of equations (1) and (2), are good predictions of future HCS behavior, depends partly upon the two parameters C_k and F_k. The first of these (the marginal resource costs) can be estimated by various accounting analyses. But the second set of parameters (the elasticities) are much harder to choose. In Devon several runs were carried out to check the accuracy of models with different parameters in reproducing known historical allocations. Table 5 gives the values used in our example.

The assistance provided to Devon was not limited to a couple of model runs like this one. Canvin et al (1978) describe in more detail how the project team worked with the local planners. In this paper, however, we concentrate on the model, and in particular on how to solve it. It is perhaps surprising that the maximization of (3) subject to (1) and (2) is not straightforward. The next section explains why.

3. SOLUTION OF THE MODEL

In purely mathematical terms the problem is to find $x_{i\ell}$ and $u_{ik\ell}$, for all i,k,ℓ, satisfying

$$\sum_{\ell} x_{i\ell} = d_i \ , \qquad \forall\ i \tag{1}$$

$$\sum_{i,\ell} x_{i\ell} u_{ik\ell} = A_k \ , \qquad \forall\ k \tag{2}$$

that maximize

$$Z(\underline{x},\underline{u}) = \sum_{i,k,\ell} x_{i\ell} h_{ik\ell}(u_{ik\ell}) \ , \tag{3}$$

where

$$h_{ik\ell}(u) = \frac{C_k U_{ik\ell}}{\beta_k}\left[\left(\frac{u}{U_{ik\ell}}\right)^{\beta_k} - 1\right] \ . \tag{4}$$

There are various possible approaches, of which the most elementary would be direct numerical search. We can, however, make more use of the forms of equations (1) - (4). We note, for example, that equations (1) - (3) are linear in $x_{i\ell}$, and that if $u_{ik\ell}$ were known for all i,k,ℓ the problem would be a simple linear program (LP). Unfortunately, the coefficient terms in equations (2) and (3) are functions of the unknown variables $u_{ik\ell}$. But in both equations we can make a piecewise linear approximation such as

$$\tilde{Z}(\underline{x},\underline{u}) = \sum_{ik\ell}\left\{\begin{matrix} {}_1x_{i\ell}h_{ik\ell}(0.1) + {}_2x_{i\ell}h_{ik\ell}(0.2) \\ + \ \ldots\ldots\ldots \ + {}_{10}x_{i\ell}h_{ik\ell}(1.0)\end{matrix}\right\} \ , \tag{5}$$

by introducing programming variables ${}_jx_{ik\ell}$, $j = 1,\ldots, 10$, that satisfy

$${}_jx_{i\ell} = \left\{\begin{matrix} x_i & : & j = \bar{j} \ , & \forall\ i,\ell \\ 0 & : & j \neq \bar{j} \ , & \forall\ i,\ell \end{matrix}\right. \ .$$

In theory, LP techniques can then be used. In practice, the approach requires a computer program or LP package with special features.

This analysis might suggest that difficulties arise because of nonlinearity in equations (2) and (3). In fact, these nonlinearities can be handled using Lagrange multipliers. Doing this, we shall reveal a problem of NDO.

We formulate the dual problem

$$\min_{\underline{\lambda}} \Phi(\underline{\lambda})$$

where $\Phi(\underline{\lambda})$ is the solution to an internal problem

$$\Phi(\underline{\lambda}) = \max_{\substack{\underline{x} \geq 0 \\ \sum \underline{x} = \underline{d}}} \max_{\underline{u} \geq 0} L(\underline{x},\underline{u},\underline{\lambda})$$

$$= L(\underline{x}^*,\underline{u}^*,\underline{\lambda}) \quad ,$$

in which * denotes the optimal value or function, and

$$L(\underline{x},\underline{u},\underline{\lambda}) = \sum_{i,k,\ell} x_{i\ell} h_{ik\ell}(u_{ik\ell}) + \sum_k \lambda_k (A_k - \sum_{i,\ell} x_{i\ell} u_{ik\ell}) \quad , \tag{6}$$

is the result of adjoining the constraint of equation (2) to the function of equation (3) with Lagrange multipliers λ_k, $k = 1,2,\ldots,6$. We now have three embedded problems which we can take in turn, and under certain conditions (proved in the Appendix) the solution to this dual problem also solves the original problem.

The first, innermost problem is easy to solve. Find $\underline{u}(\underline{x},\underline{\lambda})$ so as to

$$\max_{\underline{u} \geq 0} L(\underline{x},\underline{u},\underline{\lambda}) \quad .$$

Setting $\frac{\partial L}{\partial \underline{u}}$ equal to zero gives

$$u^*_{ik\ell}(\underline{x},\underline{\lambda}) = r_k U_{ik\ell} \quad , \tag{7}$$

provided that $x_{i\ell} \neq 0$, where

$$r_k = (\lambda_k/C_k)^{-F_k} \quad , \tag{8}$$

are "reduction factors" which, when applied to the ideal resource levels $U_{ik\ell}$, give the actual resource levels. In the model, the r_k do not vary across patient categories or modes of care, and the balance between the reduction factors for different resource types is controlled largely by the elasticities F_k. The result defined by equation (7) is always positive and therefore satisfies the constraint on $\underline{u}$. The result of the maximization is

$$L(\underline{x},\underline{u}^*,\underline{\lambda}) = b + \sum_{i\ell} c_{i\ell} x_{i\ell} \quad , \tag{9}$$

where

$$b = \sum_k \lambda_k A_k$$

$$c_{i\ell} = \sum_k \frac{C_k U_{ik\ell}}{\beta_k} \left[\frac{1}{F_k} \left(\frac{\lambda k}{C_k} \right)^{1-F_k} - 1 \right] \quad ,$$

Strictly, $c_{i\ell}$ is determined only when $x_{i\ell} \neq 0$. However, when $x_{i\ell} = 0$, the corresponding terms in equation (9) are zero anyway.

The second problem is also easy to solve. Find $\underline{x}(\underline{\lambda})$ so as to

$$\max_{\substack{\underline{x} \geq 0 \\ \sum \underline{x} = \underline{d}}} L(\underline{x},\underline{u}^*,\underline{\lambda}) \quad .$$

This is a very simple LP, for which the solution can be found by inspecting $c_{i\ell}$.

$$x^*_{i\ell} = \begin{cases} d_i & : \quad \ell = \bar{\ell} \\ 0 & : \quad \ell \neq \bar{\ell} \end{cases} \tag{10}$$

where

$$c_{i\bar{\ell}} = \max_{\ell}\{c_{i\ell}\} \quad .$$

Strictly, this unique solution for $\underline{x}^*$ exists only when there is a single mode in each category with maximum $c_{i\ell}$. Typically, however, categories have more than one such mode, and in such circumstances a unique solution for $\underline{x}^*$ cannot be found until $\underline{\lambda}^*$ is determined. Nevertheless, the result of the maximization is unaffected, being equal to

$$\begin{aligned} \Phi(\underline{\lambda}) &= L(\underline{x}^*, \underline{u}^*, \underline{\lambda}) \\ &= \sum_k A_k \lambda_k + \sum_i c_{i\bar{\ell}} d_i \end{aligned} \quad . \tag{11}$$

There remains the third problem of choosing $\underline{\lambda}$ so as to

$$\min_{\underline{\lambda} \geq 0} \Phi(\underline{\lambda}) \quad . \tag{12}$$

The difficulty here is that small continuous changes in $\underline{\lambda}$, while causing small continuous changes in $\underline{c}$, can cause large and discontinuous changes in the LP solution for $\underline{x}^*$. Because of this, $\Phi(\underline{\lambda})$ is a nonsmooth function of $\underline{\lambda}$. Loosely speaking, it has "corners" like the graph in Figure 1. Solution methods which ignore this fact may fail, especially when the solution lies on a corner. What is the meaning of a solution for $\underline{\lambda}$ on a "corner" of $\Phi(\lambda)$? It means that more than one mode in each category has maximum $c_{i\ell}$, and patients in these categories are divided between two or more modes of care. It is these mixed-mode solutions, in which there is no unique solution for $\underline{x}^*$ until $\underline{\lambda}^*$ is found, that complicate the analysis. However, once the optimal $\underline{\lambda}^*$ is found, the values of $\underline{u}^*$ are also fixed and the determination of which modes are active in each category is a straightforward LP problem.

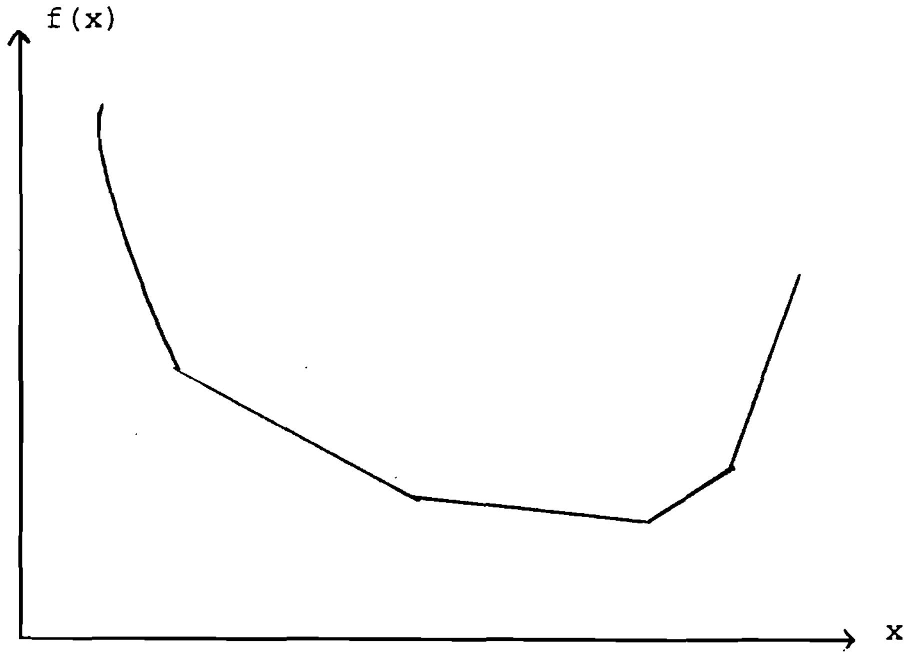

Figure 1. An example of a nonsmooth function.

The results derived above show that the problem formulated at the beginning of this section can be solved by the procedure depicted in Figure 2. The two innermost problems are solved by using equations (7), (10) and (12) to determine $\Phi(\lambda)$ for a particular choice of $\underline{\lambda}$. The way in which an NDO algorithm can be used to find the value of $\underline{\lambda}$ that minimizes $\Phi(\underline{\lambda})$ is described in the next section.

4. SOLUTION OF THE EXAMPLE

In the previous section, we showed how a solution to the example given in Section 2 can be easily found, once we have a procedure for finding the $\underline{\lambda}$ which solves the NDO problem of

$$\min_{\underline{\lambda}\geq 0} \Phi(\underline{\lambda}) \quad .$$

Such procedures are extensions of the procedures used for differentiable optimization. Where the latter use a gradient, NDO procedures use a subgradient defined as

$$g_{\underline{\lambda}} = \partial\Phi(\underline{\lambda})/\partial\underline{\lambda} \quad .$$

Unlike the gradient, the subgradient is not unique. There is a set of supporting hyperplanes at any point of nondifferentiability, and this is one of the additional features that NDO procedures must handle.

Another obstacle to be overcome is that the subgradient does not generally tend to zero as the solution is approached. This makes it difficult to identify the neighborhood of the optimum. Furthermore, the direction of $-g_{\lambda}$ is not generally one in which $\Phi(\underline{\lambda})$ decreases, and a single member of a subgradient set provides very scant information about descent directions.

Methods to solve NDO problems began to appear in the mid-sixties, and Balinski and Wolfe (1975) can be recommended as a source of references and basic ideas. Devon's problem was solved using the method described in Nurminski and Zhelikhovski (1974) to regulate the step size in a generalized

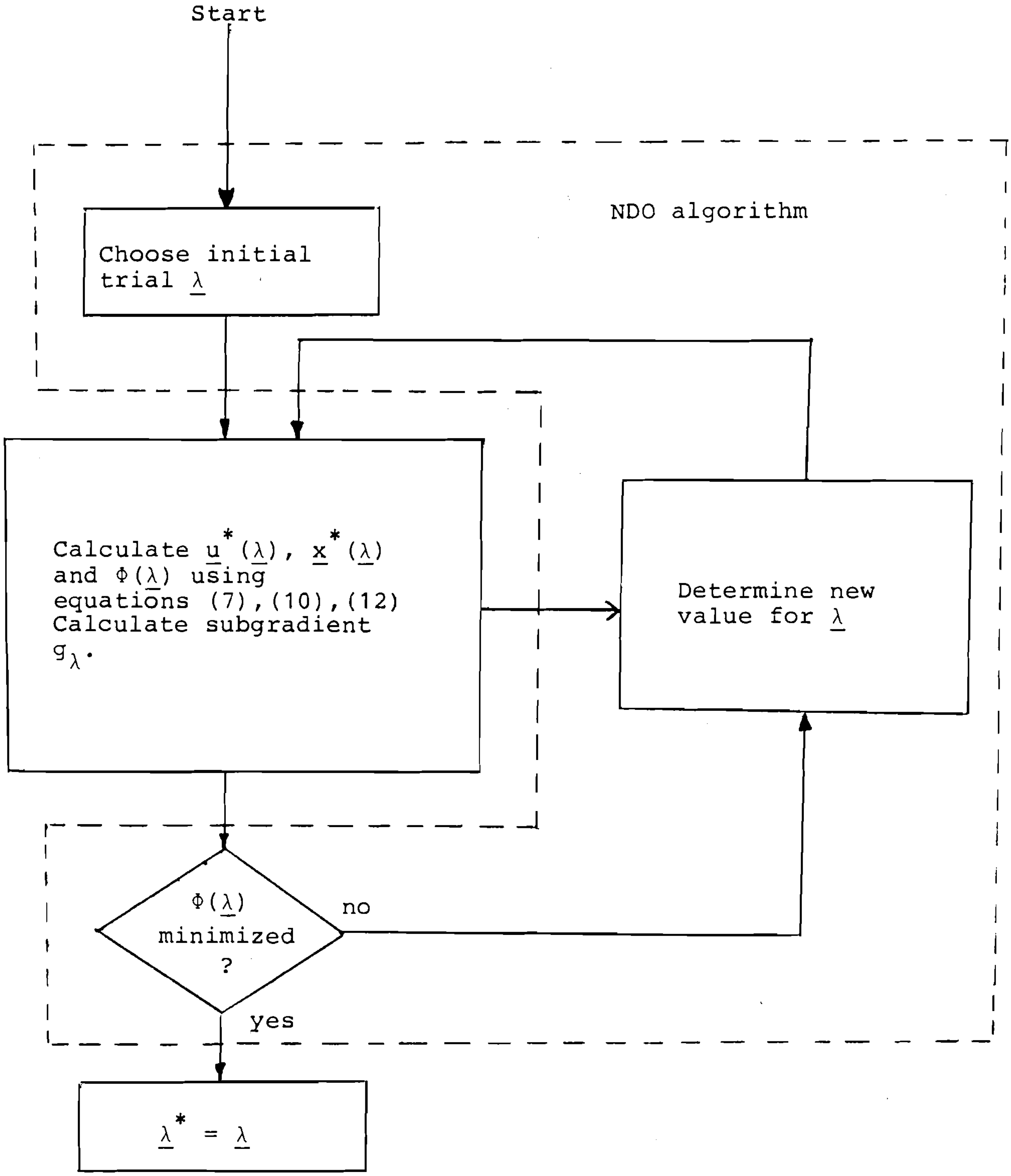

Figure 2. Solution procedure.

classical descent procedure. Although the original problem has (17*4) + (17*4*6) = 476 variables, the dual problem has only 6 variables, and hence has negligible storage requirements. The second author wrote a computer program with about 50 FORTRAN statements, which makes repeated calls of a subroutine written by the first author to calculate Φ and its subgradient. The results tabulated below were found by the IIASA PDP11/70 mini-computer with UNIX time-sharing operating system. This system makes convergence times difficult to assess. Subsequently, however, the computations were confirmed with the commercially available NDO solution routines developed by Lemarechal (1978). It took 0.5 CPU second to get a solution with machine precision on an IBM 370/168.

The same example was also solved by the third author using the piecewise linear approximation described at the beginning of Section 2. The computer package which was used (called SCICONIC) had the necessary separable programming facility with associated matrix generation and report writing. Starting from the solution to a similar problem, the central part of the SCICONIC solution (the solution of the linearized problem as a large LP) took 64 iterations and 1.7 CPU seconds; slightly longer than the NDO solution. A solution from "scratch" might have taken up to twice as long.

Table 6 gives the results obtained both by NDO and by piece-wise linear approximation. Although the second method neither uses nor calculates the Lagrange multipliers $\underline{\lambda}$ used by the first method, the reduction factors $\underline{r}$ of equation (8) are calculated by both methods and provide an equivalent comparison. We see that they are practically identical, the small differences ($<1\%$) probably being due to rounding. We conclude that both methods reached the same solution. The allocations of patients to modes of care are identical in 12 modes and different in the remaining 5. These differences arise not from the different solution methods but from the discontinuous nature of the solution for $\underline{x}^*$ as a function of $\underline{\lambda}$. Because this particular example was part of a hypothetical scenario, a direct validation of these predictions for Devon is impossible. However, similar runs have shown that the reduction factors can be quite accurately

Table 6. Solutions to Devon example by NDO and linear approximation.

Reduction Factors (r_k)

		Solution via NDO	Solutions via linear approximation
Resource types (k)	1	0.741	0.745
	2	0.451	0.453
	3	0.373	0.374
	4	0.652	0.653
	5	0.536	0.536
	6	0.257	0.257

Allocation of patients to modes ($x_{i\ell}$)

		Solution via NDO				Solution via linear approximations			
		Modes (ℓ)				Modes (ℓ)			
		1	2	3	4	1	2	3	4
Patient categories (i)	1	43				43			
	2			38				38	
	3	277		49		233	93		
	4		7	83		63	27		
	5			200				200	
	6	20		871				891	
	7				703				703
	8		184				184		
	9		818				818		
	10	43	8			51			
	11	196	2			169	29		
	12			132				132	
	13			777				777	
	14			918				918	
	15			3410				3410	
	16		339				339		
	17			2667				2667	

predicted (Coverdale and Negrine, 1978), although the actual use of different modes of care is usually more homogeneous than predicted by the model. Canvin et al (1978) give some more results for Devon. The extreme modal allocations can be regarded as optimistic predictions of reallocations within the HCS, giving reduction factors that are slightly higher than would be obtained in practice. When historical factors seem likely to prevent this, appropriate constraints can be easily applied in the model and incorporated in either method of solution.

5. CONCLUSION

The example analyzed here is interesting because it tests alternative ways to solve a practical example. Although the NDO solution was faster, it had none of the diagnostic or presentational printouts available from the SCICONIC solution, being written primarily to see how a different method would solve the example. On the other hand, the programming of a full-scale solution program to use NDO would appear to be straightforward. Because the main burden of computing falls on the subroutine that solves the internal problem (and not on the NDO routines) there is more room to extend the scope of the model wherever this might be necessary. Provided that modifications to the model do not damage the duality results exploited in the solution, the small NDO routines can remain unchanged.

From the point of view of resource allocation modelling, the new analysis of this example makes plain what solving the model actually means, and helps discussions about whether the right model is being solved. Within the framework of strategic planning in Devon, the results of Table 5 indicate how current levels of care are likely to change, and suggests what pattern of model allocation will follow if the many agents in the HCS act (or can be encouraged to act) so as to maximize levels of care.

REFERENCES

Balinski, M.L., and P. Wolfe (eds.) (1975) Nondifferentiable Optimization, Mathematical Programming Study 3. Amsterdam: North-Holland.

Canvin, R., J. Hamson, J. Lyons, and J.C. Russell (1978) Balance of care in Devon: joint strategic planning of health and social services at AHA and county level. Health and Social Services Journal 88(4604):C17-C20.

Coverdale, I.L., and S.M. Negrine (1978) The balance of care project: modelling the allocation of health and personal social services. Journal of the Operational Research Society 29(11):1043-1054.

Hughes, D.J., and A.P. Wierzbicki (1979) DRAM: a Model of Health Care Resource Allocation. RR-79-00. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Lemarechal, C. (1978) Nonsmooth Optimization and Descent Methods. RR-78-4. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Lemarechal, C., and R. Mifflin (eds.) (1978) Nonsmooth Optimization. Proceedings of anIIASA Workshop, March 28-April 8, 1977, Pergamon Press.

McDonald, A.G., C.C. Cuddeford, and E.M.L. Beale (1974) Mathematical models of the balance of care. British Medical Bulletin 30(3):262-270.

Nurminski, E.A., and A.A. Zhelikhovski (1974) Investigation of one regulating step. Cybernetics 10(6):1027-1031.

Shigan, E.N., D.J. Hughes, and P.I. Kitsul (1979) Health Care Systems Modelling at IIASA: A Status Report. SR-79-4. Laxenburg, Austria: International Institute for Applied Systems Analysis.

APPENDIX : Duality Results

Section 3 shows how the original problem is troublesome because it is nonconvex with respect to $\underline{x},\underline{u}$. It might create difficulty in finding optimal values of primal variables x,u when optimal values of dual variables are given. Generally, in nonconvex cases there is a duality gap between primal and dual problems, and in these cases direct use of a duality approach is hindered it it is possible at all. Fortunately, this does not occur in our case due to the convexity of the sets of primal variables which maximize the Langrangian for the dual variables given. These sets $S(\lambda)$:

$$S(\lambda) = \{(x_\lambda, u_\lambda) : \Phi(\lambda) = L(x_\lambda, u_\lambda)\} \quad ,$$

consist in fact of a unique u and a set of x which are solutions of the obviously convex LP problem discussed in Section 3.

For optimal λ^* which solves problem (12) we can show that the set $S(\lambda^*)$ contains the optimal primal variables x^* (there is no problem with u^* due to its uniqueness). In fact, so far as λ^* is optimal, there is a zero subgradient of the function $\Phi(\lambda)$ (11) at the point λ^*. Correspondingly there are points x^j, u^j, $j = 1,2,\ldots, \in S(\lambda^*)$ and nonnegative weights α_j such that

$$\sum_j \alpha_j = 1$$

and

$$\sum_j \alpha_j \partial L(\underline{x}^j, \underline{u}^j, \underline{\lambda})/\partial \underline{\lambda} \,\Big|_{\underline{\lambda} = \underline{\lambda}^*} =$$

$$= \sum_j \alpha_j \Big(\sum_{i\ell} x_{i\ell}{}^j u^*_{ik\ell}(\underline{\lambda}^*) - A_k \Big) \tag{13}$$

$$= \sum_{i\ell} u^*_{ik\ell}(\underline{\lambda}^*) x^*_{i\ell} - A_k = 0 \quad ,$$

where

$$x^*_{i\ell} = \sum \alpha_j x^j_{i\ell} \quad ,$$

also lies within the convex set $S^*(\underline{\lambda}^*)$. But, with equation (13), this solution satisfies constraint (2). This guarantees that the dual solution $\underline{x}^*(\underline{\lambda}^*), \underline{u}^*(\underline{\lambda}^*)$ also solves the original problem.

RELATED IIASA PUBLICATIONS IN HEALTH CARE SYSTEMS AND IN NONDIFFERENTIALBE OPTIMIZATION

Mifflin, R. (1976) Semismooth and Semiconvex Functions in Constrained Optimization. RR-76-021.

Mifflin, R. (1977) An Algorithm for Constrained Optimization with Semismooth Functions. RR-77-003.

Propoi, A.I. (1977) On the Theory of Max-Min. RM-77-013.

Gibbs, R.J. (1977) Health Care Resource Allocation Models--A Critical Review. RM-77-53.

Lemarechal, C. (1978) Nonsmooth Optimization and Descent Methods. RR-78-004.

Gibbs, R.J. (1978) The IIASA Health Care Resource Allocation Sub-model: Mark 1. RR-78-8.

Gibbs, R.J. (1978) A Disaggregated Health Care Resource Allocation Model. RM-78-1.

Mifflin, R., C. Lemarechal, and M.A. Keyzer (1978) Computing Economic Equilibria through Non-smooth Optimization. RM-78-013.

Hughes, D.J. (1978) The IIASA Health Care Resource Allocation Sub-model: Mark 2--The Allocation of Many Different Resources. RM-78-50.

Mifflin, R., and C. Lemarechal (1978) Nonsmooth Optimization. CP-78-002.

Gibbs, R.J. (1978) Computer Programs for the IIASA Health Care Resource Allocation Sub-model, Mark 1--A User's Guide. WP-78-15.

Hughes, D.J. (1978) The IIASA Health Care Resource Allocation Sub-model: Formulation of DRAM Mark 3. WP-78-46.

Nurminski, E.A. (1978) Non-Differentiable Optimization with ε-Subgradient Methods. WP-78-55.

Nurminski, E.A. (1978) On ε-Differentiable Mappings and their Application in Nondifferentiable Optimization. WP-78-58.

Ermoliev, Yu.M. (1978) Methods of Nondifferentiable and Stochastic Optimization and their Applications. WP-78-62.

Wierzbicki, A. (1978) Lagrangian Functions and Nondifferentiable Optimization. WP-78-63.

Hughes, D.J. (1978) The IIASA Health Care Resource Allocation Sub-model: Estimation of Parameters. RM-78-67.

Nurminski, E.A. (1979) Conceptual Newton Method for Solving Multivalued Inclusions: Scalar Case. WP-79-50.