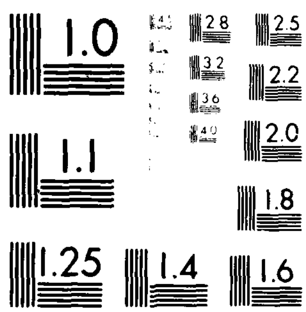
AD-A098 406 FOREIGN TECHNOLOGY DIV WRIGHT-PATTERSON AFB OH F/G 12/1
ELEMENTS OF DYNAMIC PROGRAMMING,(U)
FEB 81 Y S VENTTSEL'
UNCLASSIFIED FTD-ID(RS)T-1515-80 NL

1 OF 4
AD
A098406



MICROCOPY RESOLUTION TEST CHART

FTD-ID(RS)T-1515-80

2

FOREIGN TECHNOLOGY DIVISION

ELEMENTS OF DYNAMIC PROGRAMMING

by

Ye. S. Venttsel'

Approved for public release;
distribution unlimited.

FTD- ID(RS)T-1515-80

UNEDITED MACHINE TRANSLATION

FTD-ID(RS)T-1515-80 2 February 1981

MICROFICHE NR: FTD-81-C-000040

ELEMENTS OF DYNAMIC PROGRAMMING

By: Ye. S. Venttsel'

English pages: 311

Source: Elementy Dinamicheskogo Programmirovaniya, ~~Publishing House~~ "Nauka", Moscow, 1964 [redacted] 1-174.

Country of origin: USSR
This document is a machine translation
REquester: FTD/TQF
Approved for public release; distribution unlimited.

THIS TRANSLATION IS A RENDITION OF THE ORIGINAL FOREIGN TEXT WITHOUT ANY ANALYTICAL OR EDITORIAL COMMENT. STATEMENTS OR THEORIES ADVOCATED OR IMPLIED ARE THOSE OF THE SOURCE AND DO NOT NECESSARILY REFLECT THE POSITION OR OPINION OF THE FOREIGN TECHNOLOGY DIVISION.	PREPARED BY: TRANSLATION DIVISION FOREIGN TECHNOLOGY DIVISION WP-AFB, OHIO.

TABLE OF CONTENTS

U. S. BOARD ON GEOGRAPHIC NAMES TRANSLITERATION SYSTEM

Block	Italic	Transliteration	Block	Italic	Transliteration
А а	*А а*	A, a	Р р	*Р р*	R, r
Б б	*Б б*	B, b	С с	*С с*	S, s
В в	*В в*	V, v	Т т	*Т т*	T, t
Г г	*Г г*	G, g	У у	*У у*	U, u
Д д	*Д д*	D, d	Ф ф	*Ф ф*	F, f
Е е	*Е е*	Ye, ye; E, e*	Х х	*Х х*	Kh, kh
Ж ж	*Ж ж*	Zh, zh	Ц ц	*Ц ц*	Ts, ts
З з	*З з*	Z, z	Ч ч	*Ч ч*	Ch, ch
И и	*И и*	I, i	Ш ш	*Ш ш*	Sh, sh
Й й	*Й й*	Y, y	Щ щ	*Щ щ*	Shch, shch
К к	*К к*	K, k	Ъ ъ	*Ъ ъ*	"
Л л	*Л л*	L, l	Ы ы	*Ы ы*	Y, y
М м	*М м*	M, m	Ь ь	*Ь ь*	'
Н н	*Н н*	N, n	Э э	*Э э*	E, e
О о	*О о*	O, o	Ю ю	*Ю ю*	Yu, yu
П п	*П п*	P, p	Я я	*Я я*	Ya, ya

*ye initially, after vowels, and after ъ, ь; e elsewhere.
When written as ё in Russian, transliterate as yё or ё.

RUSSIAN AND ENGLISH TRIGONOMETRIC FUNCTIONS

Russian	English	Russian	English	Russian	English
sin	sin	sh	sinh	arc sh	$\sinh^{-1}$
cos	cos	ch	cosh	arc ch	$\cosh^{-1}$
tg	tan	th	tanh	arc th	$\tanh^{-1}$
ctg	cot	cth	coth	arc cth	$\coth^{-1}$
sec	sec	sch	sech	arc sch	sech^{-1}
cosec	csc	csch	csch	arc csch	csch^{-1}

Russian	English
rot	curl
lg	log

ELEMENTS OF DYNAMIC PROGRAMMING.

Ye. S. Venttsel'

Page 2.

Annotation.

Dynamic programming - recently emergent and intensely developing section of mathematics, which gives methods for the solution of important practical problems. Discussion deals with planning/gliding of production or other processes when control by them is realized by a multistage path in view of their complexity. To such tasks can be attributed, for example, the selection of the most advantageous profile/airfoil for laying out the railway line (decomposed into the series/row of sections), the selection of the best sizes/dimensions of the steps/stages of multistage rocket and many others.

In this book for the first time in the Soviet literature is done the attempt to accessibly present basic ideas and methods of dynamic programming.

The book is of interest for the wide circle of the workers of science and production, and also for all persons, developers of contemporary science interesting.

Page 3.

Preface.

In the book is given the elementary presentation of the method of dynamic programming which is considered as the general method of the construction of optimum control by different types of physical systems. The book is intended for the engineers, the economists and the scientific workers of the different specialties, which are occupied by questions of planning/gliding, and also by the selection of the rational parameters of technical devices/equipment. The author did not place to himself by the task of giving strict and consecutive presentation of the mathematical aspect of method, and he attempted to make him clear and available for the wide circle of practical workers, who do not have special mathematical formation/education and interested mainly in the direct use/application of a method to their tasks interesting. This target determined by itself the style of the presentation accepted: the book barely contains strict proofs; the explanation of the principles of method is conducted with the support to the multiple practical tasks and the examples of which many are led to the concrete/specific/actual numerical result. Tasks and examples are undertaken from the most varied regions of practice; in the presentation are stressed the general/common/total features, which make it possible to decide by their similar methods.

Page 4.

The mathematical apparatus, used in the book, is simple and nowhere it exceeds the limits of the course of advanced calculus, set forth in all VTUZ [Higher technical educational institution], and for the most part it does not require even this and is reduced to the simple arithmetic and algebraic operations. However, for the conscicus mastering of material is required the known stress/voltage of thought. By somewhat unusual ones for the inexperienced reader can seem the used with the presentation general formulas; however, the sense of these formulas and figuring in them designations is in detail explained in the text. For the understanding of two latter/last paragraphs (§§ 15 and 16) is required the acquaintance with the elementary concepts of the probability theory.

The dismantled/selected at the book specific problems intended are selected very simple that the cumbersome calculations would not shield the entity of method. In practice, as a rule, it is necessary to be encountered with the more complex problems for solving which it is necessary to draw contemporary electronic computational engineering. Keeping in mind the need for the composition of machine algorithms, the author for the elongation/extent of the entire book uses the standard logic circuit of the construction of the process of step by step optimization and the standard sequence of formulas, which facilitates programming for EVTsM [electronic digital computer].

The author consciously bounded himself by the examination only of the discretized problems of dynamic programming with a finite number of steps/pitches m and left aside the limiting cases, which correspond m→∞ and to the unlimited decrease of the length of step/pitch. However, the distinct mastering of the idea of method on the elementary tasks can substantially facilitate to reader, who desires to obtain more intimate knowledge, further study of object/subject on the more solid managements/manuals.

Ye. Venttsel'.

Page 5.

§ 1. Task of dynamic programming.

Dynamic programming (or, otherwise, "dynamic programming") is the special mathematical apparatus, which permits implemention of optimum planning/gliding of the controlled processes. Under "those controlled" are understood the processes to course of which we can to one or the other degree affect.

Widely-known close attention, given by contemporary science to questions of planning/gliding in all regions of human activity. Most general problem of optimum (best) planning/gliding is placed as follows.

Let be assumed to the realization certain action or the series of the actions (is shorter, "operation"), which pursues the specific target. It does request itself: how it is necessary to organize (to plan) operation so that it would be most efficient, i.e., in the best way satisfied the stated before it requirements?

So that stated problem of optimum planning/gliding would gain quantitative, mathematical character, it is necessary to introduce into the examination certain numerical criterion W, by which we will characterize quality, success, efficiency of operation.

Value W, depending on the character of the decided task, can be chosen by different methods. For example, during planning/gliding of the activity of the system of industrial enterprises as criterion W can be (according to the circumstances) selected the total yearly volume of production or the net revenue; the criterion of the efficiency of the work of transport system can be, for example, general/common/total goods freight turnover or the mean cost/value of the shipment of the ton of load.

Page 3.

The criterion of the efficiency of the bombing raid can be, for example, average/mean area of the caused destruction either an average number of affected objects, or the cost/value of the reducing works which it is necessary to fulfill opponent.

Generally the criterion of efficiency in each specific case is chosen on the basis of the purposeful directionality of operation and task of research (what element of control is optimized and for which).

Task of rational planning/gliding - to select this method of organizing this system of operations in order to become maximum (or the minimum) some criterion W. If as the criterion is undertaken such value whose increase to us is profitable (for example, income from the group of enterprises), then it they attempt to become maximum. If, on the contrary, value W it is profitable to reduce, then it they attempt to become the minimum. It is obvious, the task of the minimization of criterion easily is reduced to the task of maximization (for example, by sign change of criterion). Therefore subsequently in the examination of the tasks of planning/gliding in the general/common/total setting we will frequently speak simply about the "maximization" of criterion W.

Let us give now quantitative, mathematical posing of general problem of optimum planning/gliding.

There is certain physical system S whose state in the course of time varies. Process is controlled, i.e., we have the capability to a certain degree to affect its course, choosing at its discretion the another control U. With the process is connected certain value (or criterion) W, which depends on the used control. It is necessary to select this control U so that the value W would become maximum.

Contemporary mathematical science has available the whole arsenal of the methods, which make it possible to solve the task of optimum control. Among them special position occupies the method of dynamic programming. The specific character of this method in the fact that for finding the optimum control the planned/glide operation is divided into the series/row of consecutive "steps/pitches" or "stages". Respectively the very process of planning/gliding becomes "multistage" and is developed consecutively/serially, from one stage to the next, moreover each time is optimized control only at one step/pitch.

Page 7.

Some operations naturally fall into the stages; in others this articulation is necessary to build-in artificially.

Let us consider an example "logically multistage" operation. Let be planned/glided the activity of certain system of the industrial enterprises

$$\Pi_1, \Pi_2, \ldots, \Pi_k$$

for certain period of time T, which consists of m of economic years (Fig. 1.1).

In the beginning of period T for the development of the system of enterprises are selected some basic means K; furthermore, the functioning enterprises give some income, which is realized at the end of each year in the form of pure/clean gain.

In the beginning of each fiscal year (i.e. at moments/torques $t_1, t_2, \ldots, t_i, \ldots, t_m$) is produced financing the entire system of enterprises, moreover for each of them is selected some share of the means, available at this time at the disposal of the planning/gliding organ/control.

Let us designate $x_i^{(j)}$ the sum, separated in the beginning of the i year in the share of enterprise Π_j.

Is raised the question: how it is necessary to distribute on the enterprises initial capital K and incomes entering so that toward the end of the period of planning/gliding T total income from the entire system of enterprises would be maximum?

The formulated task is the typical task of multistage planning/gliding.

Let us look, are such they can be approaches to the solution of this problem.

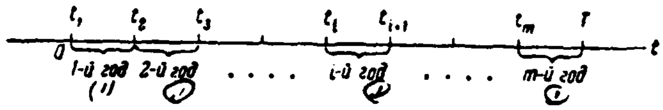


Fig. 1.1.

Key: (1). y year.

Page 3.

Let us assume that the distribution of means on by i-th the step/pitch of operation is carried out, i.e., we selected the specific control U_i:

$$U_i=(x_i^{(1)},\ x_i^{(2)},\ \ldots,\ x_i^{(k)}). \qquad (1.1)$$

Formula (1.1) is read as follows: control U_i on by i-th step/pitch lies in the fact that we isolated to enterprise P_1 of means $x_i^{(1)}$, to enterprise P_2 - means $x_i^{(2)}$, and so forth.

Using widely used terminology, control U_i it is possible to visualize as vector in a k- graduated space whose components are equal to $x_i^{(1)},\ x_i^{(2)},\ \ldots,\ x_i^{(k)}$.

Let us consider entire set of controls (allocated resources)

$$U_1,\ U_2,\ \ldots,\ U_m \qquad (1.2)$$

at m steps/pitches of cperat... as m of vectors in a k- dimensional space. The criterion of ..., W of multistage operation, as which we did select total ... of years, does depend on the entire set of controls (1.2):

$$W = W(U_1, U_2, \ldots, U_m). \quad (1.3)$$

Is raised the question: as to select control at each step/pitch, i.e., as to distribute means, so that the value W would take maximum value?

The posed by us based on specific example problem easily can be generalized.

Let be planned/glided the operation, which falls to m of consecutive steps/pitches or stages. In the beginning of each (i-th) stage it is necessary in a ... manner to select available parameters

$$x_i^{(1)}, x_i^{(2)}, \ldots$$

set of which

$$U_i = (x_i^{(1)}, x_i^{(2)}, \ldots)$$

forms control in the i stage.

Page 9.

As it is necessary to select the set of the controls

$$U_1, U_2, \ldots, U_m$$

so that certain value W, which depends on it, would become the maximum:

$$W = W(U_1, U_2, \ldots, U_m) = \max?$$

The method of dynamic programming makes it possible to produce this optimum planning/glading step by step, optimizing in each stage only one step/pitch.

Generally speaking, this approach to the determination of the optimum solution is not the only possible. The task of planning the multistage processes in the principle admits another solution - direct, with which all steps/pitches are joined into one.

Actually/really, let us consider criterion W as function from the elements of control at each step/pitch:

$$W = W(x_1^{(1)}, x_1^{(2)}, \ldots; x_2^{(1)}, x_2^{(2)} \ldots; \ldots; x_m^{(1)}, x_m^{(2)}, \ldots) \tag{1.4}$$

This function of many arguments can be traced to the maximum, as such, without the necessary distribution of the elements/cells of control "on the steps/pitches". For this it is necessary to find this value part of arguments $x_i^{(j)}$ $(i = 1, 2, \ldots, m;\ j = 1, 2, \ldots)$, with which function

(1.4) reaches maximum.

It would seem, by what simpler? It is necessary to use for the determination of maximum the classical method: to differentiate function W of all arguments, to equate derivatives zero and to solve the obtained system of equations:

$$\frac{\partial W}{\partial x_1^{(1)}} = 0, \ \frac{\partial W}{\partial x_1^{(2)}} = 0, \ \ldots;$$
$$\frac{\partial W}{\partial x_2^{(1)}} = 0, \ \frac{\partial W}{\partial x_2^{(2)}} = 0, \ \ldots; \ \frac{\partial W}{\partial x_m^{(1)}} = 0, \ \frac{\partial W}{\partial x_m^{(2)}} = 0, \ \ldots \quad (1.5)$$

However, this simplicity is illusory.

First, when steps/pitches much, this method becomes very bulky. The task of solving the system of equations (1.5) in the simplest cases only proves to be easily solvable. As a rule, it is very complicated, and frequently it is easier directly" grope" the maximum of function (1.4), than to solve system of equations (1.5).

Page 10.

Furthermore, the method indicated does not completely guarantee the determination of the solution ! Let us recall that by itself rotation/access of derivative into zero does not ensure the maximum of function, and is always required further checking. Moreover, this method does not give the possibility to find maximum, if it lies/rests not inside, but on the boundary of the region of the possible values of the arguments (for example, see Fig. 1.2: the absolute maximum of function $\phi(x)$ is reached not at points x_1, x_2, x_3, where the derivative is equal to zero, but at end-point x' of the region, in which is preset the function). So that even in those rare cases when system of equations (1.5) can be solved, finding absolute maximum requires the whole system of checkings, the more complicated, the greater the arguments in function.

Finally, it is necessary to supplement that in the series/row of practical tasks function w generally cannot be differentiated; for example, when the elements of control $x_1^{(1)}, x_1^{(2)}, \ldots; x_2^{(1)}, x_2^{(2)}, \ldots; x_m^{(1)}, x_m^{(2)}, \ldots$ are the not continuously not changing, but discrete/digital values.

All these circumstances lead to the fact that the use/application of classical methods of analysis (or the calculus of variations) to the solution of the majority of the tasks of planning/gliding proves to be ineffective: it reduces initially stated problem of finding the maximum to such secondary tasks which prove to be not simpler than the initial, but often also it is more complicated.

At the same time the solution of such many problems can be substantially simplified, if we develop the process of planning/gliding step by step, i.e., by the method of dynamic programming. The idea of method in the fact that finding the maximum of the function of many variable/alternating is substituted by the repeated finding of the maximum of function of one or small number of variable/alternating.

What in this case are applied methods, it will be evidently from the following presentation.

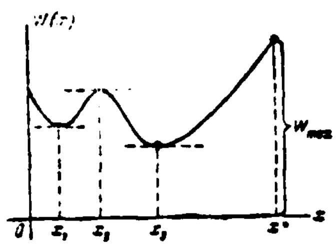


Fig. 1.2.

Page 11.

§ 2. Principle of the step by step construction of optimum control.

Thus, dynamic programming is step by step planning/gliding of the multistage process, during which in each stage is optimized only one step/pitch.

At first glance it can seem that formulated idea is sufficiently trivial. Actually/really, that the here odd: if it is difficult to optimize control immediately for the elongation/extent of entire operation, then to decompose it into the series/row of consecutive steps/pitches and to optimize separately each step/pitch. Not so whether?

Completely not thus ! The principle of dynamic programming does not in any way assume that, choosing control at one single step/pitch, it is possible to forget about all others. On the contrary, control at each step/pitch must be chosen taking into account all its aftereffects in the future. Dynamic programming - these are planning/gliding farsighted, taking into account prospect. This not short-sighted planning "blindly" for one step ("come what may provided was now good"). On the contrary, control at each step/pitch is chosen on the basis of the interests of operation as a whole.

Let us illustrate the principle of "farsighted" planning/gliding based on examples.

Let, for example, be planned/glided the work of the group of heterogeneous industrial enterprises for the period of time m of years and final task is obtaining the maximum capacity of the production of certain class C of consumers' goods.

In the beginning of period is a specific supply of means of the production (machines, equipment), with the help of which it is possible to begin the production of goods of this class.

By "step/pitch" or "stage" of the process of planning/gliding is fiscal year. Let for us be in prospect the selection of the solution for the purchase of raw material, machines and the distribution of means according to the enterprises to the first year. During the "short-sighted" step by step planning/gliding we would make the decision: to put a maximum quantity of means into the purchase of raw material and to release existing machines at full power, whereas approaching the maximum capacity of the production of class C toward the end of the first year.

Page 12.

To what it can give this planning/gliding? To the rapid wear of machine park and, as a result, to the fact that on the second year the production will fall.

During the farsighted planning/gliding, on the contrary, will be provided the actions, which ensure filling machine park in proportion to its wear. Taking into account such investments the capacity of the production of the basic goods C in the first year will be less than it could be, but will be provided the possibility of expanding the production during the subsequent years.

Let us take another example. The process of planning/gliding in the checkered game also will fall into the single steps/pitches (courses). Let us assume that the figures are conditionally evaluated by one or the other number of glasses with respect to their importance; taking figure, we win this number of glasses, and giving up - we play back.

Reasonably whether it will, thinking over chess match on several steps/pitches forward, always approach at each step/pitch to win a maximum number of glasses? It is obvious, no. This, for example, solution, as "to endow figure", never can be profitably from a narrow point of view of only one course, but it can be profitably from the point of view of match as a whole.

So is matter, also, in any region of practice. Planning/gliding multistage operation, we must choose control at each step/pitch, on the basis not of the narrow interests of precisely this step/pitch, but of the wider interests of operation as a whole, and hardly ever these two points of view coincide.

But how to construct this control? We already formulated the general rule: in the process of step by step management planning at each step/pitch must be accepted taking into account the future. However, from this rule there is an exception/elimination. Among all steps/pitches there is one, which can be planned/glided simply, without the "caution to the future". What this is step/pitch? Is obvious, the latter. This latter/last step/pitch, single of all, can be planned/glided so that it as such would yield the greatest profit.

After planning optimally this latter/last step/pitch, it is possible to it "to attach" next-to-last, to this in turn, of prepenultimate, etc.

Page 13.

Therefore the process of dynamic programming is always turned/run up in the opposite of the time direction: not from the beginning toward the end, but from the end/lead at the beginning. First of all is planned/glided latter/last step/pitch. but as it to plan, if we do not know how did end next-to-last? It is obvious, it is necessary to do different assumptions about that how ended next-to-last step/pitch, and for each of them to select control on the latter.

This optimum control, selected under the specified condition about that how ended the previous step/pitch, we will call conditional optimum control.

The principle of dynamic programming requires determination at each step/pitch of conditional optimum control for any of the possible issues of the preceding step.

Let us demonstrate the diagram of this procedure. Let be planned/glided a m- step operation, and it is unknown how ended (m-1) -th step. Let us do about this a series/row of "hypotheses" or "assumptions". These hypotheses we will designate:

$$S^{(1)}_{m-1},\ S^{(2)}_{m-1},\ \ldots,\ S^{(i)}_{m-1},\ \ldots \qquad (2.1)$$

We will be specified, that by letter $S_{m-1}^{(l)}$ is not compulsorily designated one number: this can be and the group of the numbers, which characterize issue (m-1) -th of step but can be and the simply qualitative state of that physical system, in which proceeds the controlled process.

Let us find for each of assumptions (2.1) conditional optimum control on the latter/last (by the m-th) step/pitch. This will be that of all possible controls U_m at which attains a maximally possible value gained at the latter/last step/pitch.

Let us assume that for each of assumptions (2.1) conditional optimum control U_m^* on the latter/last step/pitch is found:

$$U_m^*(S_{m-1}^{(1)});\quad U_m^*(S_{m-1}^{(2)});\ \ldots;\quad U_m^*(S_{m-1}^{(l)});\ \ldots \quad (2.2)$$

This means that the latter/last step/pitch is planned for any issue of next-to-last.

Page 14.

Let us switch over to planning/gliding of following from the end/lead, next-to-last step/pitch. Let us again do a series/row of hypotheses about that how enmd prepenultimate ((m-2) -th) the step:

$$S_{m-2}^{(1)},\ S_{m-2}^{(2)},\ \ldots,\ S_{m-2}^{(k)},\ \ldots \quad (2.3)$$

Let us raise the question: how it is necessary to choose for each of these hypotheses conditional optimum control at (m-1) -th the step?

It is obvious, it must be chosen so that it, together with the already selected control at the latter/last step/pitch, would ensure the maximum value of criterion w at two latter/last steps/pitches.

In other words, for each of hypotheses (2.3) it is necessary to find this conditional optimum control on (m-1) -th step $U_{m-1}^{*}(S_{m-2})$, so that it, in conjunction with already obtained conditional optimum control $U_{m}^{*}(S_{m-1})$, would give a maximally possible prize at two latter/last steps/pitches.

It is obvious, toward (m-1)-and to step/pitch thus accurately it can be connected (m-2) -th and so forth up to the quite latter/last (from the end/lead) 1st step/pitch from which the process begins.

The first step/pitch, in contrast to all others, is planned/glided somewhat otherwise. Since we usually know, from what begins the process, then for us is no longer required to make hypotheses about state in which we begin the first step/pitch. This state to us is known. Therefore, taking into account that all subsequent steps/pitches (the 2nd, the 3rd, etc.) are already planned (conditionally), to us it remains simply to plan the first step/pitch so that it would be optimum taking into account all controls, already accepted in the best way by all subsequent steps/pitches.

The principle, placed as the basis of the construction of such solution (to seek the always optimum continuation of process relative to that state, which is achieved/reached at the given moment) they frequently call the principle of the optimum character (see [1]).

The general/common/total explanation of the method of the construction of the optimum solution by the method of dynamic programming which was given in the present paragraph, in view of quite its generality can seem by incomprehensible. Therefore the following paragraphs (§§ 3-5) we will dedicate to the solution of the specific problems on which let us try to give to him more intelligible interpretation. Subsequently, into § 6, we again will return to the general/common/total formulation of the problem which will prove to be clearer against the background of the already dismantled/selected specific examples.

Page 15.

§3. Task about the gain of altitude and velocity.

One of the simplest tasks, solved by the method of dynamic programming, is the task about the optimum climb and velocity flight vehicle. From this task we begin presentation of practical procedures of dynamic programming, moreover for the purpose of systematic clarity the conditions of task will be to the extreme simplified.

Task consists of the following. The aircraft (or another aircraft), which is found on height H_0 and which has velocity V_0, must be raised to base altitude $H_{кон}$, and its velocity is led to preset value $V_{кон}$. Is known the fuel consumption, required for lifting the apparatus from any height H_1 on any another $H_2>H_1$ at the constant velocity V; is known also the fuel consumption, required for an increase in the velocity from any value of V_1 to any other $V_2>V_1$ at the constant/invariable height H.

It is necessary to find the optimum climb and velocity during which general/common/total fuel consumption will be minimum.

The solution we will construct is follows. For simplicity let us assume that entire process of the gain of altitude and velocity is divided into the series/row of consecutive steps/pitches (stages), and for each step/pitch aircraft increases only height or only velocity.

We will be depict the state of aircraft with the help of the point on certain plane VOH, where the abscissa represents the velocity of aircraft V, and ordinate - its height H (Fig. 3.1).

The process of the displacement of point S, which represents the state of aircraft, from the initial state S_0 into final $S_{кон}$ will be depicted on plane VOH as certain stepped broken line. This line (trajectory of the motion of point S on plane VOH) will characterize control of the process of the gain of altitude and velocity.

Page 16.

It is obvious, there are many possible controls - many trajectories on which it is possible to translate point S from S_0 in $S_{кон}$. Of these all trajectories it is necessary to select that on which the selected criterion W (fuel consumption) will be minimum.

In order to construct the solution by the method of dynamic programming, let us divide height $H_{кон}-H_0$, which must be assembled on aircraft, on n_1 of the equal parts (for example, to six, see Fig. 3.2), and velocity $V_{кон}-V_0$, which it is necessary to gather, on n_2 of equal parts (for example, to eight). Let us divide the process of the gain of altitude and velocity into the single steps/pitches and we will consider that for one step/pitch the aircraft can either increase height by value

$$\Delta H=\frac{H_{кон}-H_0}{n_1},$$

or velocity - to value

$$\Delta V=\frac{V_{кон}-V_0}{n_2}.$$

A number of parts n_1, n_2 into which are divided intervals $H_{кон}-H_0$, $V_{кон}-V_0$, fundamental value does not have and can be selected on the basis of the requirements for the accuracy of the solution of problem. Pair of numbers n_1, n_2 determines by itself the total number of steps/pitches m of the multistage process of the gain of altitude and velocity

$$m=n_1+n_2.$$

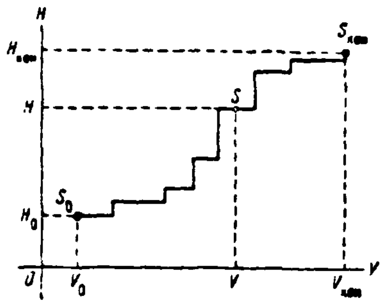


Fig. 3.1.

Page 17.

Of our case (Fig. 3.2) any trajectory will consist of fourteen steps/pitches:

$$m = 6 + 8 = 14$$

(as, for example, each of two trajectories, noted by arrows/pointers in Fig. 3.2). Being moved from S_0 in $S_{кон}$ point S can move only over the horizontal and vertical segments.

Let us register on each of these segments (Fig. 3.3) the corresponding to it fuel consumption in some arbitrary units [1].

FOOTNOTE [1]. The digits, given in Fig. 3.3, are selected from the systematic considerations and nothing in common with the real fuel consumption they have. ENDFOOTNOTE.

Any trajectory, which translates point S from S_0 in $S_{кон}$, is connected with the specific fuel consumption. For example, the trajectory, depicted as arrows/pointers in Fig. 3.3, gives the fuel consumption, equal to $W=12+11+10+8+11+8+10+10+13+15+20+9+12+14=163$ (conditional, units).

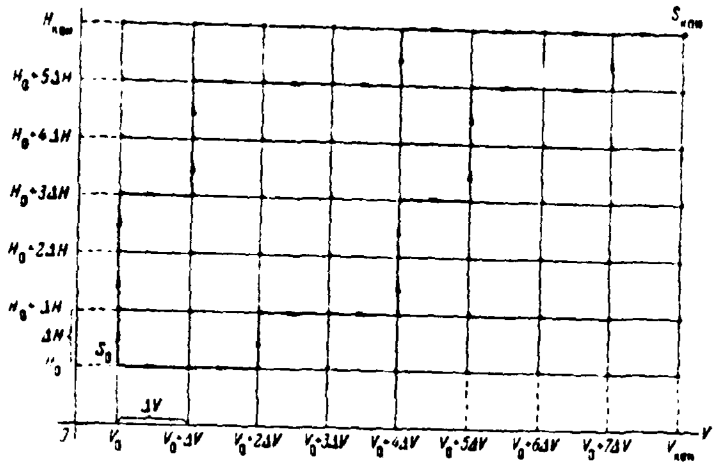


Fig. 1.2.

Page 18.

It is obvious, there is a very large number of different trajectories, which translate S of S_0 in $S_{кон}$ and to each of them corresponds its fuel consumption W. We should of such all trajectories find optimum - that, on which the fuel consumption is minimal. It would be possible it goes without saying to sort out all possible trajectories and in the final analysis to find optimum, but this - very bulky path. Much more often it is possible to solve task by the method of dynamic programming on the steps/pitches.

Process consists of m=14 steps/pitches; we will optimize each step/pitch, beginning from the latter. The final state of aircraft - point $S_{\text{кон}}$ on plane VOH - to us is preset. The fourteenth step/pitch without fail must lead us into this point. Let us look, whence we can move into point $S_{\text{кон}}$ at the fourteenth step/pitch.

Let us consider the separately right upper angle of rectangular grid (Fig. 3.4) with end point $S_{\text{кон}}$. In point $S_{\text{кон}}$

6 it is possible to move of two adjacent points (B_1 and B_2), moreover of each - only in one manner, so that the selection of conditional control at the latter/last step/pitch we do not have any - it is singular. If next-to-last step/pitch led us into point B_1, then we must move over the horizontal and expend 17 units of fuel; if we into point B_2, go on the vertical line and to expend 14 unity.

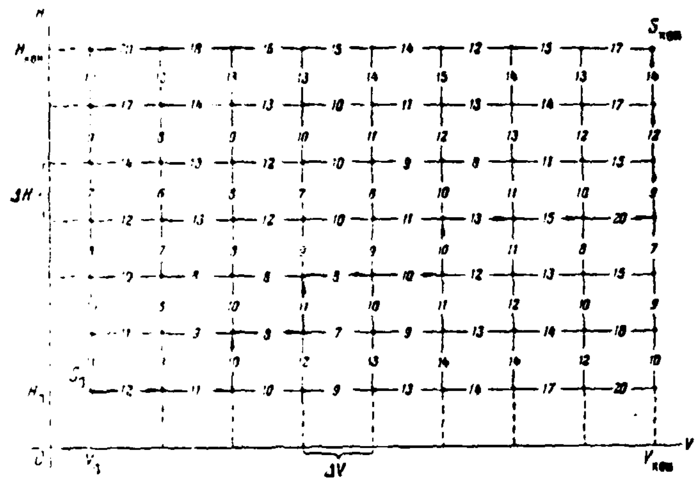


Fig. 3.3.

Page 19.

Let us register these minimum (in this case simply unavoidable) fuel consumption in the special small circles which let us supply at points B_1, B_2 (Fig. 3.5). The recording by "17" in the small circle in B_1 it indicates: "if we arrived in B_1, then the minimum fuel consumption, which translates us into point $S_{ном}$, was equal to 17 unity". Analogous sense has a recording by "14" in the small circle at point B_2. The optimum control, which leads to this expenditure/consumption, is marked in each case by the arrow/pointer, which emerges from the small circle. Gunner/rifleman indicates the direction over which we must move from this point, if as a result of our previous activity they proved to be in it.

Thus, conditional optimum control on the latter/last fourteenth step/pitch is found for any (B_1 or B_2) issue of the thirteenth step/pitch. For each of these issues it is found, furthermore, minimum fuel consumption due to which of this point it is possible to move in $S_{кон}$.

Let us switch over to planning/gliding of next-to-last (thirteenth) step/pitch. For this we should consider all possible results of prepenultimate (twelfth) step/pitch. After this step/pitch we can prove to be only in one of the points C_1, C_2, C_3 (Fig. 3.6). From each such point we must find optimum path in point $S_{кон}$ and corresponding to this path minimum fuel consumption.

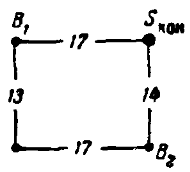


Fig. 3.4.

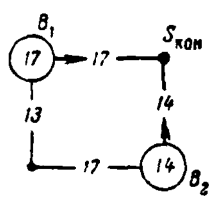


Fig. 3.5.

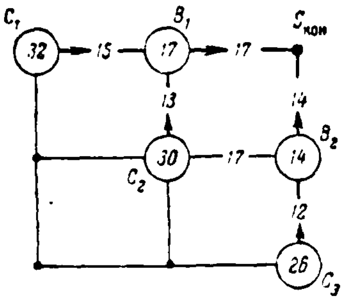


Fig. 3.6.

Page 20.

For point C_1 there is no selection: we must be moved on the horizontal and expend 15+17=32 unity of fuel. This expenditure we will register in the small circle with point C_1, and optimum (in this case single) path from point C_1 again let us mark by arrow/pointer.

For point C_2 the selection exists: from it carried it is possible to go in $S_{кон}$ through B_1 or through B_2. In the first case we will spend 13+17=30 unity of fuel; the secondly 17+14=31 unity. It means, optimum path from C_2 is vertical (let us note this by arrow/pointer), and minimum fuel consumption it is equal to 30 (this number we will register in the small circle with point C_2).

Finally, for point C_3 path into $S_{кон}$ the again single: on the vertical line; is bypassed ... into 12+14=26 unity; this value (26) we let us register in the small circle with C_3, and by arrow let us mark optimum control.

Thus, passing from one point to the next from right to left and downward (from the end/lead of the process to its beginning), it is possible for each node Fig. 3.3 to select conditional optimum control at the following step i.e., the direction, which leads in $S_{кон}$ with the minimum fuel consumption, and to register in the small circle with this point this minimum expenditure/consumption. In order to find from each point the optimum following step/pitch, it is necessary to trace two possible ... paths from this point: to the right and upward, and for each path to find the sum of fuel consumption per this step and minimum fuel consumption per the optimum continuation, already constructed from the following point, where is directed pointer tip. From two paths is chosen that, for which this sum is less (if sum they are equal, it is chosen any of the paths).

Thus, from each point Fig. 3.3 (see page 18) is conducted the arrow/pointer, which indicates optimum path from this point (optimum conditional control), and in the small circle it is entered/written the fuel consumption, reached at the optimum control, beginning from this point to the end/lead.

Sooner or later this process of the construction of conditional optimum controls is finished, after reaching starting point S_0. From this point as of any another, conducts the arrow/pointer, which indicates, where it is necessary to be moved in order to reach S_{kon} optimally. After this it is possible to construct entire optimum trajectory, being moved on the arrows/pointers, already from the beginning of process to its end/lead.

Page 21.

Fig. 3.7 shows the final result of this procedure - optimum trajectory, which leads from S_0 in $S_{кон}$ on the arrows/pointers, i.e., having from each point optimum continuation. This trajectory is noted by fatty/greasy small circles and tailcale countars. The number "139", which stands at point S_0, indicates minimum fuel consumption W*, less which it cannot be obtained in what trajectory.

Thus, stated problem is solved and optimum control of the process of the gain of altitude and velocity is found. It consists of the following:

on the first step/pitch to increase only velocity, retaining by constant/invariable height H_0, and to bring velocity to $V_0+\Delta V$;

at the second step/pitch to increase height to $H_0+\Delta H$, retaining the velocity of constant/invariable;

at the third, fourth and fifth steps/pitches to again gain speed, until it becomes equal to $V_0+4\Delta V$;

at the sixth, seventh and eighth steps/pitches to gain altitude and to bring it to $H_0+4\Delta H$;

at the ninth, tenth, eleventh and twelfth steps/pitches to again gain speed and to bring it to preset finite value $V_{кон}$.

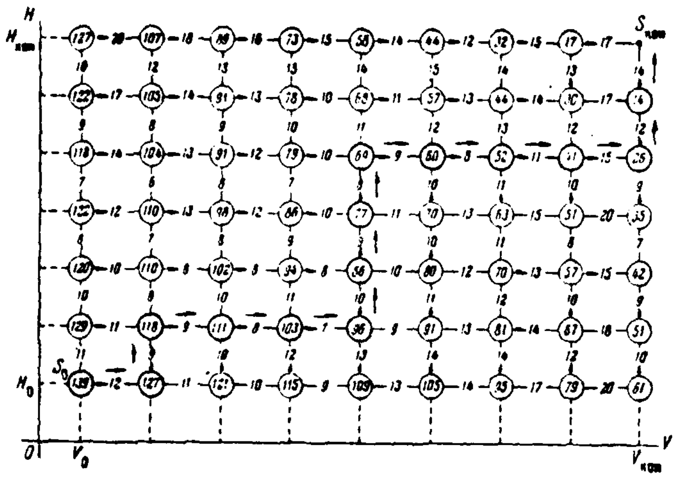


Fig. 3.7.

Page 22.

at the latter/last two steps/pitches (the thirteenth and the fourteenth) to gain altitude to preset value $H_{ком}$.

It is not difficult in a number of examples to ascertain that the obtained control is actually/really optimum, i.e., which in any other trajectory, which leads of S_0 in $S_{ком}$ the fuel consumption will be more.

Task examined here of the optimum gain of altitude and velocity is the simplest example in which they frequently demonstrate the basic idea of dynamic programming. Actually/really, in our simplified setting the problem greatly easily is solved to the end/lead with the help of the simplest methods. This is explained by the following circumstances. First, at each step/pitch for us it is necessary to choose not more than between two versions of control (" to gain altitude" or "to gain speed"). The determination of conditional optimum control at each point elementarily and is reduced to the selection of of more advantageous of these two paths. In the second place, in our task it is very simple to produce the numbering of steps/pitches, beginning from the end/lead. Actually/really, each trajectory consists of one and the same number of steps/pitches, and latter/last, naturally, proves to be that which by overcoming one (horizontal or vertical) step/stage, directly, gives into point $S_{кон}$; with next-to-last - that, after which to point $S_{кон}$ there remains only one step/pitch and, etc.

This simplified formulation of the problem does not completely correspond to reality. Actually flight vehicle can to gather (often it gathers) height and velocity simultaneously.

Let us try (furthermore in the simplified form) to pose the problem where will be provided this simultaneous set, and let us look, to what complications of methodology this will lead.

Let us return to the diagram in Fig. 3.3 and will change it so that, besides the already examined paths (upward and to the right), from each node of grid would be feasible another path along the diagonal of rectangle (simultaneous gain of altitude and velocity). Let us supply the appropriate fuel consumption along each diagonal (Fig. 3.8).

What does differ this diagram from the previous (see Fig. 3.3)? Not only the presence, except two as possible earlier controls, the still third "along the diagonal".

Page 23.

This diagram differs the less precise numbering of the steps/pitches: in the limits of each rectangle from the lower left angle into the upper right it is possible to pass both for two steps/pitches ("upward-to the right" or "to the right-upward") and for one step/pitch - along the diagonal. Therefore in the new task is unsuitable this simple principle of consecutive sorting nodes, as what we took earlier (according to a number of steps/pitches, which remained to the end/lead), and it is necessary to take some another.

Let us agree to label nodes not according to a number of steps/pitches, which remained to the end/lead, but according to the sign/criterion of any coordinate. As this coordinate it is possible to take, for example, "remainder/residue of velocity", $V_{\text{кон}} - V$ which must be "reached" for the remaining time. With this numbering of points the "latter" will be that step/pitch which will translate point S with the vertical straight line (m-1)-(m-1) (Fig. 3.8) into point $S_{\text{кон}}$ (this latter/last "step/pitch" can consist of several steps/stages); by "next-to-last ones" - that which will translate point with the straight line (m-2)-(m-2) to the straight line (m-1)-(m-1), and so forth.

Fig. 3.9 examines the sample/specimen of the optimization of process with this numbering of the steps/pitches (are shown only two latter/last steps/pitches).

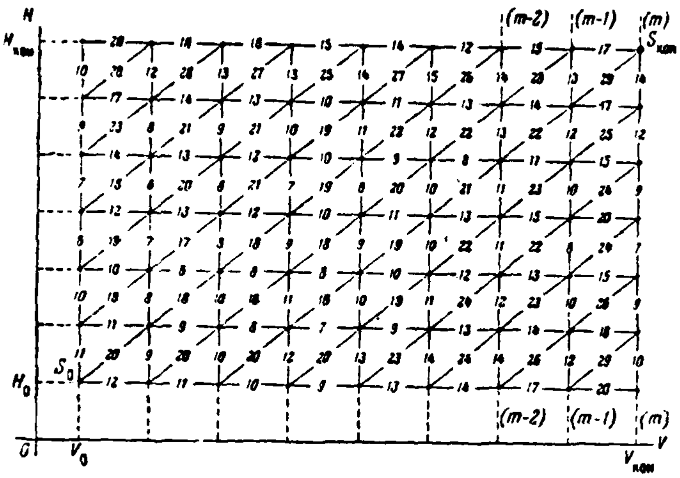


Fig. 3.8.

Page 24.

Conditional optimum control, as earlier, it is shown by arrows/pointers. Let us clarify the procedure of the construction of control.

If we proved to be at any point on the straight line (m)-(m), passing through $S_{кон}$ that the only possible (the very same and optimum) path of output into point $S_{кон}$ - but vertical line. This path is shown in Fig. 3.9 by arrows/pointers of lengthwise entire straight line (m)-(m); the corresponding fuel consumption are shown in the small circles.

Let us assume now that as a result of the process of the gain of altitude and velocity we proved to be on the straight line (m-1)-(m-1). We will sort out on this straight line all points on top downwardly. If we proved to be in the juice peak, then path into point $S_{кон}$ hence single (horizontal) is bypassed it in 17 unity of the fuel (we write/record 17 in the small circle and we place horizontal arrow/pointer). We pass to the following point - the second on top. From it to the straight line (m)-(m) - three paths. The first path - upward - to the right - is bypassed into 13+17=30 unity of fuel; the second - along the diagonal - in 29 unity; the third - to the right - upward - in 31 unity. We choose diagonal path, we mark by its arrow/pointer, and the corresponding consumption - 29 unity - we place in the small circle. For the third point we are on top again congruent/equate three paths:

upward (and further along the diagonal): 12+29=41 unity of fuel;

with respect to the diagonal (and further upward): 25+14=39 unity of fuel;

to the right (and further top): 15+26=41 unity of fuel.

We choose optimum path - along this diagonal, we note by arrow/pointer, we write/record 39 in the small circle. For the following - the fourth on top - point on the straight line (m-1)-(m-1) optimum will be the path upward and so forth.

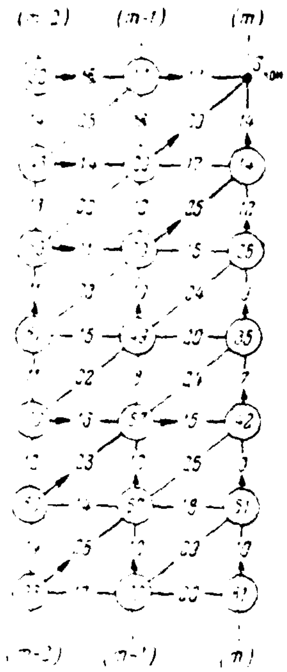


Fig. 3.9.

Page 25.

Fig. 3.10 gives the final results of the optimization of control of the gain of altitude and velocity under given conditions, indicated in Fig. 3.8. Optimum trajectory is as before isolated with fatty/greasy small circles and [illegible] counters.

Being congruent/equating the optimum trajectory, shown in Fig. 3.10, with that that is given in Fig. 3.7, we note that they are distinguished not very not strongly. As far as fuel consumption is concerned, then its values (130 and 13?) entirely differ little from each other, and both controls can be considered in effect equivalent.

In the example examined we selected the method of the ordering of steps/pitches "on the abscissa". This method is not completely necessary; it would be possible to label steps/pitches also in terms of the values of any another coordinate. As this coordinate could be with the equal success selected height H. Perhaps, in our example the most natural "ordering" coordinate would be distance from $S_{кон}$ deposited/postponed in parallel to the diagonal of basic rectangle (Fig. 3.11).

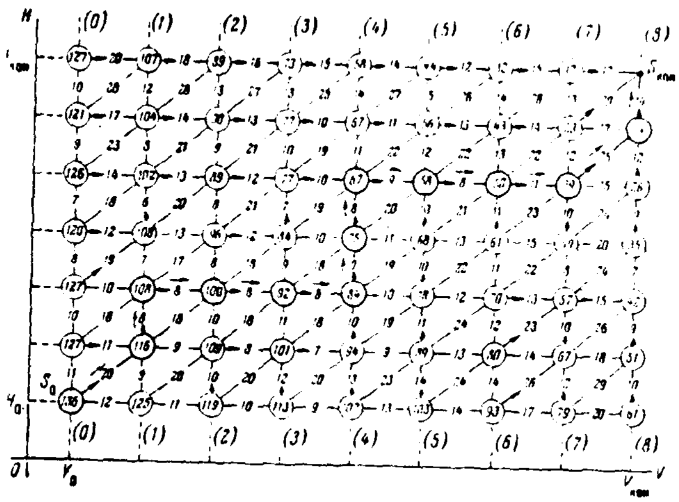


Fig. 3.1).

Page 26.

To reader one should by way of exercise find optimum trajectory according to the data of Fig. 3.8, using that method of the adjustment of steps/pitches which is demonstrated in Fig. 3.11.

In the diverse tasks of the dynamic programming where there is no natural distribution into the steps/pitches, the principle of this distribution and ordering of steps/pitches is chosen depending on convenience in the organization of computational process, taking into account to the required accuracy of the solution of problem. Is generally intuitively it is clear that with an increase in the number of steps/pitches the accuracy of the solution grows. In some tasks it proves to be possible to obtain even limited solution with $m \to \infty$; this solution can represent theoretical, and sometimes also practical interest; however, usually it is to sufficient explain the structure of optimum control in the general/common/total, rough features; in this case there is no need to strongly increase a number of steps/pitches. The same with the practical realization of control most frequently fits nevertheless to step back from the strictly optimum version which can prove to be difficultly to feasible. Therefore we will not pause on the maximum tasks of control, which appear with $m \to \infty$, but we will be bounded to the examination of discrete/digital step by step diagram. This especially makes sense, that in many tasks of the economic planning/gliding (but we will give to similar tasks considerable attention) distribution into the steps/pitches not imposed from outside, but it is logically dictated by the discrete/digital nature of the planning/gliding itself (ylane is comprised, for example, by the year, to the month, etc., and it does not vary continuously in the course of production process).

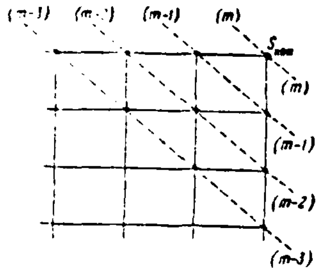


Fig. 3.11.

§ 4. Problem of the selection of the fastest path.

In the previous paragraph we considered in the simplest setting the task of the optimum gain of altitude and velocity. Here we will consider similar in the type, but nevertheless differing somewhat from it task of the selection of the fastest path of one point/item in another.

Page 27.

Task is placed in the following manner. Let it be we should reach in the machine from point/item S_0 to point/item $S_{кон}$ (Fig. 4.1). Generally there is a whole series of possible versions of path. They are comprised of the sections of ways, not equivalent or the quality. Among them there can be, for example, the sections of the first-class asphalted highway, and also the less well-organized and simply unimproved roads. Furthermore, on the path to us can be met the crossings and the passages on which the motion is detained.

Task lies in the fact that to select this path from S_0 and $S_{кон}$ which machine will pass for the minimum time.

Task at first glance is completely similar to that examined in the previous paragraph. However, it has some special features/peculiarities. In examples 9 3 we constructed the regular, rectangular grid of the nodes through which could pass trajectory.

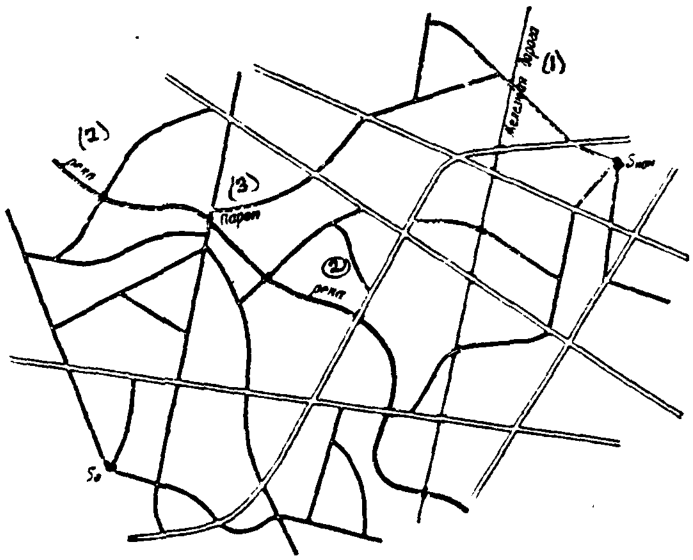


Fig. 4.1.

Key: (1). Railroad. (2). river. (3). ferry boat.

Page 28.

In the new task which we examine now, the role of this grid of nodes could play the logically noted "singular points" of the network/grid of ways - crossings and passages, on they were arranged/located too irregularly, and it is difficult "to order them on the steps/pitches". In order to approximately solve our task by the method of dynamic programming, it is possible to introduce into it artificially certain "stage-by-stage character". For example, it is possible to divide distance of D from S_0 to $S_{кон}$ into m of equal parts length ΔD=D/m (Fig. 4.2) and to consider that for each "step/pitch" of the process of displacement from S_0 in $S_{кон}$ is surmounted the m part of distance D (in the direction $S_0-S_{кон}$). In other words, each "step/pitch" is displacement with one of the lines of support, perpendicular $S_0-S_{кон}$. to the adjacent, the closer to $S_{кон}$.

Dale process to the steps/pitches thus, we, naturally, must agree that the displacement from one step/pitch to the next is allowed/assumed only in the positive direction (i.e. from S_0 to $S_{кон}$ and not conversely); in other words, after the specific step/pitch is travelled, return conversely, into the same band between two lines of support, is not allowed/assumed. This limitation occurs sufficiently to acceptable ones for the practice. Let us recall that in the task § 3 we met with even the more severe limitation: in the mode/conditions of climb and velocity was allowed/assumed the displacement over both axes only in the positive direction. In the task of the selection of the fastest path that introduced by us limited (as a result of each step/pitch to be moved only "there", but to direction $S_0-S_{кон}$. and not"back ") is less rigid, since it functions only from one step/pitch to the next, not within the step/pitch, and moreover, only cn one axis (in the case of necessity from this limitation it is possible to be freed, cn in this case the solution strongly it is complicated).

Thus, let us assume that the path from S_0 in $S_{\text{кон}}$ is decomposed into m of the steps/pitches, in each of which the machine is moved with one of the lines of support $(i)-(i)$ to the following in order

$$(i+1)-(i+1) \qquad (i=0,\ 1,\ \ldots,\ m).$$

The carried out by us lines of support intersect road net at some points.

Page 29.

For the solution of problem to us must be known the time, required for the passage of each section of path, and also delay time on each crossing (passage). In Fig. 4.2 against each route segment is written the corresponding time of passage (in the minutes), and in the small circle in each crossing (passage) - latency of machine in this point/item.

According to a quantity of lines of support in Fig. 4.2 process of the displacement of machine from S_0 in $S_{\text{кон}}$ we will share into seven steps/pitches (i.e. let us take m=7) and let us begin the construction of optimum path from the latter/last (m-th) step/pitch.

Let us plan on the line of support (m-1)-(m-1) all possible positions of machine at the moment of the termination of next-to-last (m-1) step/pitch. This will be hypotheses about the state of machine afterward (m-1) step/pitch, for each of which we must find conditional optimum control on the m step/pitch. In Fig. 4.2 these possible position are noted by small circle with the point inside. From each such position we must pave optimum (shortest on the time) path into point $S_{кон}$.

Let us consider first the first (on top) of the noted points - point A on the straight line (m-1)-(m-1). From it into point $S_{кон}$ (in the limits of the band of the m step/pitch) conducts one-single path, which occupies on the time 10+2+1+5+10+2+5=35 (minutes).

The selection of this path is conditional optimum control when the previous step/pitch led us into point A. Let us note in Fig. 4.2 this optimum path by black heavy line, and in point A let us display "flag" with registered with it digit 35.

Heavy line together with the flag they indicate the following: if, being moved from S_0 in $S_{кон}$ we some fates they proved to be at point A, then of it we must move further over the noted by black line route and on the achievement of point $S_{кон}$ of the expenditures of 35 minutes.

We pass to the following point (B) on the straight line (m-1)-(m-1). From it into point $S_{кон}$ conducts the one and only path, to which it is required 2+2+1+5+10+2+5=27 (minutes). Number 27 we also write/record on the flag next to point B.

For point C the path again single and is continued 4+2+5=11 (minutes).

From point K in $S_{кон}$ there are two paths: 3+3+4=10 (minutes) and 3+3+2+2+6=16 (minutes); of them the first - fastest; we note by its heavy line and we write/record minimum time (10) on the flag in point K.

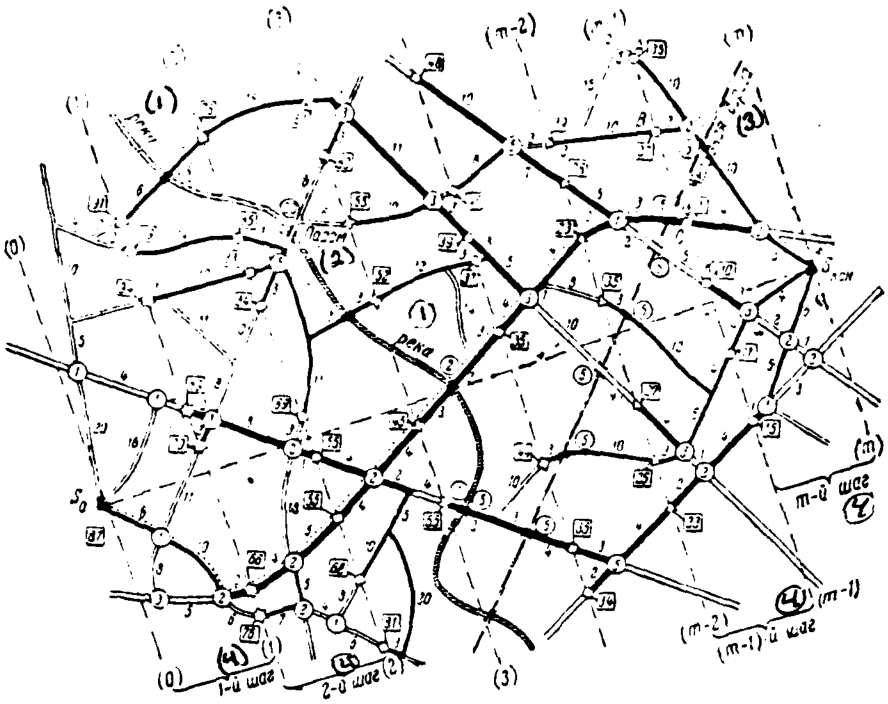


Fig. 4.2.
Key: (1). river. (2). ferry road. (3). Railroad. (4). m step/pitch.

Continuing thus, we find for each point on the straight line (m-1)-(m-1) the optimum continuation of path - conditional optimum control at the m step/pitch.

After this is carried out, we pass toward planning/gliding of (m-1) step/pitch. Hypotheses about that, where can be located the machine of afterward penpenultimate (m-2) step/pitch, they are noted by triangles on the straight line (m-2)-(m-2).

For each of the noted points we must find conditional optimum control, i.e., this path with the straight line (m-2)-(m-2) to the straight line (m-1)-(m-1), which, together with the already optimized latter/last step/pitch, gives the possibility to achieve S_{km} for the minimum time. In order to find this conditional optimum control, we must for each point on the straight line (m-2)-(m-2) sort out all possible transition modes to the straight line (m-1)-(m-1) and time, required to this transition, sum with the minimum time of latter/last step/pitch, registered with the flag. From all possible paths is chosen that, for which this total time is minimal: path is noted by black line, and time is written/recorded on the flag in the corresponding point.

As a result of the chain/network of such constructions, being moved step by step with one line of support to another, we finally will reach starting point S_0. For it we will determine optimum path to the straight line (1)-(1) and let us register the appropriate minimum time (87 minutes) on the flag at point S_0. Thus, all data for the construction of optimum path there are, since for each of the planned points (whatever fates we in it not they proved to be) is known the optimum continuation of path. In order to construct optimum path from S_0 in $S_{кон}$ it is necessary simply to be moved on the sections of ways, noted by heavy lines. In Fig. 4.2 optimum trajectory from S_0 in $S_{кон}$ is noted by heavy line with the dotted line.

Thus, stated problem about the selection of the fastest path between two preset points/items is solved.

Apropos of this task it is possible to express several considerations, which concern the selection of a number of steps/pitches during the construction of the solution by the method of dynamic programming.

Page 32.

It at first glance seems that so that the solution would be more simply, it would be desirable have a little less steps/pitches. However, this not entirely thus. The larger/coarser the step/pitch, the more difficult it is to find the optimum solution on this step/pitch, the more there is the versions of the displacement with the straight line to the straight line. In the extreme case, if we considered only one step/pitch ($m=1$), before us would arise the initial task of sorting all possible paths from S_0 in $S_{кон}$ in her entire complexity.

It does follow from this that in our specific problem it was necessary to still increase a number of steps/pitches, to do them, for example, not 7, but 20?

Also no ! An increase in the number of steps/pitches beyond some reasonable limits would only complicate the procedure of the construction of the solution.

In the fact that the selected by us number of steps/pitches (m=7) is sufficiently reasonable, is possible to be convinced on the fact that for us was nowhere necessary to sort out a large number of versions of transition with the straight line to the straight line - these versions proved to be one, two, are rare three, and to find among them optimum was not too not difficult. If we strongly increased a number of steps/pitches, i.e., is excessive refined the sections of transition, then in the overwhelming majority of the cases with the straight line to the straight line would conduct one and only path, and no optimization it would be. As the final result we would construct the same optimum trajectory, on by more complicated calculations.

§ 5. Continuous task of plotting of optimum route.

In § 4 was solved the task of plotting of optimum route from point/item S_0 to point/item $S_{кон}$ when both points/items were connected by some network/grid of ways and path can run only on one of the ways of this discrete/digital grid.

In practice can be met another situation - when finished road net there does not exist, but direction of motion from each point on the plane can be chosen arbitrarily, for example in the limits of some angular sector θ (Fig. 5.1). In this case for each point A on plane xOy is known the velocity of displacement from this point over any ray/beam AA' within the limits of sector θ.

Page 33.

Task lies in the fact that to find such trajectory U*, which combines S_0 and $S_{кон}$, along which point S would pass from S_0 in $S_{кон}$ within the short time.

Let us plan the diagram of the solution of this problem by the method of dynamic programming. For simplicity let us assume that the sector θ is symmetrical relative to line AB, parallel to the axis of abscissas, and that $\theta<180°$ (latter is necessary in order to exclude the lisplacements, "reverse" to the lirection of the axis of abscissas).

Let us decompose distance $S_{кон} - S_0$ into m of equal parts, and the process of overcoming this distance - to m of the steps/pitches, each of which is transition with one of line of support, parallel to axis ordinates, to another, adjacent (Fig. 5.2).

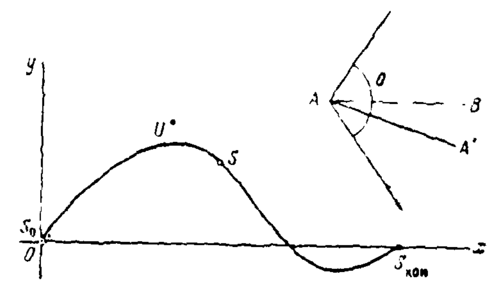


Fig. 5.1.

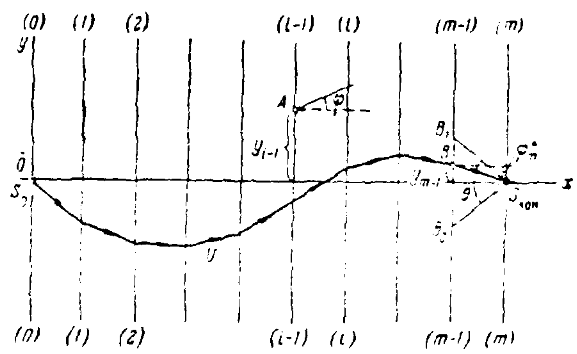


Fig. 5.2.

Page 34.

If we take a number of steps/pitches m sufficiently large, then it is possible to assure that at each step/pitch the trajectory phase is straight-line. Task is reduced for each point A on the line of support (i-1)-(i-1) to determine the optimum angle φ (in the limits of sector θ), at which must pass point A optimum trajectory, i.e., that out with which we must move from A in order to achieve $S_{кон}$ in the minimum time. If we the position of point A on the straight line (i-1)-(i-1) determine by its ordinate y_{i-1}, the conditional optimum control at the i step/pitch will be motion from point A at angle φ_i^* toward the axis of the abscissas where φ_i^* depends on y_{i-1}:

$$U_i^*(y_{i-1}) = \varphi_i^*(y_{i-1}).$$

We will plan/glide the process of displacement, as always, from the latter/last (m-th) step/pitch. Let us assume that as a result of (m-1) step/pitch we proved to be at certain point B on the straight line (m-1)-(m-1) (Fig. 5.2). Where we must move further in order to prove to be at point $S_{кон}$? It is obvious, on straight line $BS_{кон}$. But direction of this motion not for all positions of point B on the straight line (m-1)-(m-1) it is found within the permissible limits. In order to construct segment B_1B_2 of the possible positions of point B, whence it is possible, during our limitations, to arrive in $S_{кон}$, obviously, it is necessary to construct from point $S_{кон}$ the "inverted" sector θ; its borders will cut off on the straight line (m-1)-(m-1) segment B_1B_2.

Thus, for each point B cutting off B_1B_2 is found conditional optimum control - displacement over straight line $BS_{кон}$ at angle φ_m^* to the axis of abscissas, which depends on ordinate y_{m-1} of point B:

$$\varphi_m^* = \varphi_m^*(y_{m-1}).$$

Knowing the velocity of displacement from point B over this direction, we can find the minimum time of the execution of the latter/last step/pitch:

$$T^*_m = T^*_m(y_{m-1}).$$

Thus, for any point B on the straight line (m-1)-(m-1) conditional optimum control and corresponding to it conditional minimum time of the m step/pitch can be determined.

Page 35.

Let us assign several values of ordinate y_{m-1}:

$$y^{(1)}_{m-1},\quad y^{(2)}_{m-1},\ \ldots \tag{5.1}$$

and for each of them let us find conditional optimum control and conditional minimum time:

$$\varphi^*_m(y^{(1)}_{m-1});\quad \varphi^*_m(y^{(2)}_{m-1});\ \ldots$$
$$T^*_m(y^{(1)}_{m-1});\quad T^*_m(y^{(2)}_{m-1});\ \ldots$$

If points $y^{(1)}_{m-1}, y^{(2)}_{m-1}, \ldots$ are set on cutting off B_1B_2 sufficiently frequently, then it is possible to consider that the conditional optimum control and conditional minimum time are found for any value y_{m-1}.

Let us switch over to planning/gliding of next-to-last (m-1) step/pitch (Fig. 5.3). Section C_1C_2 of the possible positions of point S as a result of (m-2) step/pitch is also determined by the "inverted" sector θ; in order it to construct, it is necessary to continue straight lines $S_{кон}B_1$ and $S_{кон}B_2$ before the intersection with the straight line (m-2)-(m-2). Let us place in section C_1C_2 the series/row of reference points and for each of them let us find optimum path in $S_{кон}$. For point C_1 this path is clear: it goes on straight line $C_1S_{кон}$. Let us lead this path by heavy line and it is computed corresponding to it full/total/complete time $T^*_{m-1, m}$, expended for the execution of two latter/last steps/pitches.

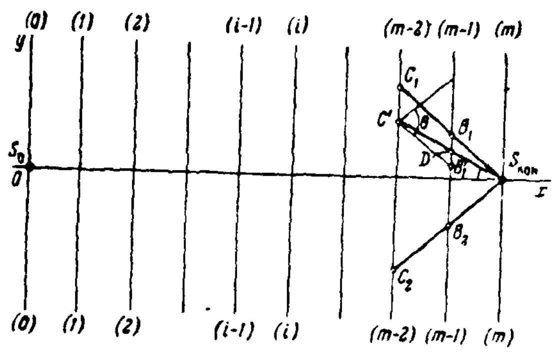


Fig. 5.3.

Page 36.

This time is equal to the sum of two times: time t^{*}_{m-1} of the displacement over cutting off C_1B_1 and of time T^{*}_{m} of displacement from B_1 in $S_{кон}$ (but it was already calculated on the previous step/pitch). Analogously optimum path from C_2 in $S_{кон}$ goes on straight line $C_2S_{кон}$.

Let us take on cutting off C_1C_2 any internal point C'. For this point path to the straight line (m-1)-(m-1) no longer single. Actually/really, after constructing with point C' the sector of possible directions θ, we see that within the limits of this sector it is possible to select any rectilinear path, which leads from C' into one of the points cutting off $B_1B'_1$. By what from these paths to select? It is obvious, that each C'D, for which the total time, which goes to both latter/last steps/pitches (C'D and $DS_{кон}$). is minimal. Let us designate this minimum time $T^{*}_{m-1,m}$ and let us note that it depends on ordinate y_{m-2} of point C'. Taking whole range of different positions of point C' on cutting off C_1C_2 with ordinates $y^{(1)}_{m-2}, y^{(2)}_{m-2}, \ldots$ let us find for each of them optimum control φ^{*}_{m-1} and minimum time $T^{*}_{m-1,m}$ of the achievement of point $S_{кон}$:

$$\varphi^{*}_{m-1}(y^{(1)}_{m-2}); \quad \varphi^{*}_{m-1}(y^{(2)}_{m-2}); \ldots$$
$$T^{*}_{m-1,m}(y^{(1)}_{m-2}); \quad T^{*}_{m-1,m}(y^{(2)}_{m-2}); \ldots$$

After this is done, we pass toward planning/gliding of (m-2) step/pitch, etc.

As a result of the chain/network of such constructions for each point on any of the lines of support will be explained the conditional optimum control (is found the angle φ*, at which it must pass optimum trajectory) and is determined corresponding to this control minimum time of the achievement of point $S_{кон}$.

After the process of optimization is led to point S_0, is constructed (already from the beginning toward the end) entire optimum trajectory, which from each point goes at optimum angle φ*.

Fig. 5.4 shows the result of the construction of optimum control by the method of dynamic programming. Optimum trajectory is noted by heavy line with the dotted line.

Page 37.

Let us note in conclusion that the described methodology of the construction of optimum trajectory completely does not depend on that, what precisely value is minimized - be it time T of displacement from S_0 into $S_{кон}$, either the fuel consumption R or of costs/values of passage C, or even any criterion, selected depending on the character of the decided practical task. For example, with the packing of railway line it is possible to prefer that route which leads to the smallest expenditures or the smallest volume of earthwork. When selecting of missile trajectory can prove to be necessary minimize the launching weight or maximize velocity, etc.

Let us note also that the sector of possible directions θ, which we for simplicity considered constant/invariable, can vary depending on the number of step/pitch or, in general, from the coordinates of starting point A. In some practical tasks it occurs, that direction the velocities at the initial moment or into the final (or into both, etc.) are preset previously and vary cannot. This is equivalent to so that the sector of possible directions θ in the vicinity of these points is degenerated into one straight line.

Reference lines (i)-(i), by which we shared process into the stages, completely must not be the straight lines, parallel to one of the axes; them are chosen, on the basis of convenience in the construction of the solution.

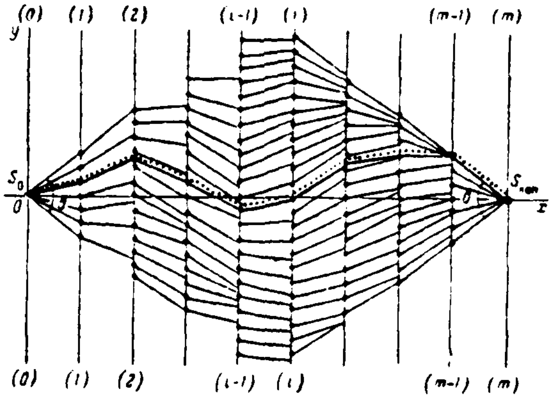


Fig. 5.4.

Page 38.

If, for example, the task more conveniently deciding not in the Cartesian, but in the polar coordinate system, line (i)-(i) they can be, depending on type the tasks, are selected in the form of light beam, which proceed from pole 0 (Fig. 5.5) or in the form of the concentric circumferences (Fig. 5.6). For example, Fig. 5.7 shows the schematic of planning/gliding the output of rocket from point S_0 on the earth's surface to the given point $S_{кон}$ of outer space, carried out in the polar coordinate system. The polar coordinates of launching point S_0 are preset and equal to (R_1, φ). The conditions of the verticality of start narrow the sector of possible directions θ at the first step/pitch to the straight line; the conditions for the preset direction of final velocity $v_{кон}$ set the same limitations on the latter/last step/pitch; at the intermediate steps/pitches of the limitations, superimposed in the sector of possible directions, they are derived/concluded from the physical considerations (for example, from the maximum permissible transverse transfers of rocket).

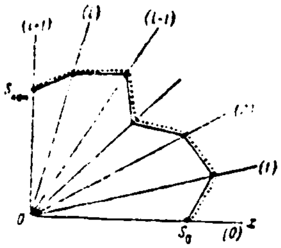


Fig. 5.5.

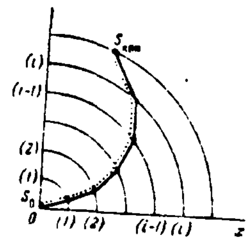


Fig. 5.6.

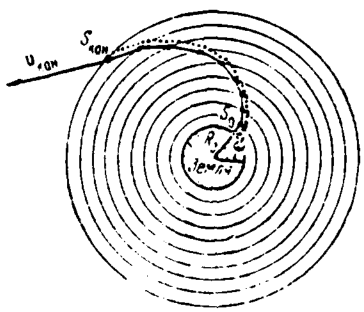


Fig. 5.7.

Page 39.

§ 6. General/common/total formulation of the problem of dynamic programming. Interpretation of control in the phase space.

After is examined the series/row of the specific problems of dynamic programming, let us give the general/common/total formulation of such problems and will formulate in general form the principles of their solution. In this paragraph (and in the following after it § 7) for reader, familiar only with the elements/cells of advanced calculus, it is necessary to clash with an not entirely customary for it recording of formulas and a somewhat unusual terminology. However, let us emphasize that the conscious mastering of precisely these paragraphs is very substantial for the understanding of the method: without this general/common/total approach will be difficult to see in the following presentation anything larger than the set of diverse examples.

Let us consider following general problem.

There is certain physical system S, which in the course of time can vary its state. We can manage this process, i.e., in this or some other way to affect the state of system, to translate it of one state into another. This system S we will call the controlled system, and action, with the help of which we affect the behavior of system, by control.

With the process of changing the state of system S is connected some of our interest, which is expressed numerically with the help of criterion W, and it is necessary to organize process so that this criterion would become maximum (minimum) [1].

FOOTNOTE [1]. Subsequently for the brevity we will speak only about the maximization of criterion, implying that the "maximum" in any event can be substituted to the "minimum". ENDFOOTNOTE.

Let us designate our control (i.e. entire system of actions with the help of which we we affect the state of system S) of one letter U. Criterion W depends on this control; this dependence we will register in the form of the formula

$$W = W(U). \tag{6.1}$$

Page 40.

It is necessary to find such control U* ("optimum control"), during which criterion W reaches the maximum:

$$W^* = \max_U \{W(U)\}. \tag{6.2}$$

Recording $\max_{U}$ is read: "maximum on U", and formula (6.2) indicates: W* is maximum from the values which take criteria W during all possible controls U.

However, the problem of the optimization of control is not completely yet posed. Usually, upon the formulation of such problems must be taken into consideration some conditions, superimposed on the initial state of system S_0 and final state $S_{кон}$. In the simplest cases both these states are completely preset (as, for example, in § 4, when it was necessary to translate truck from point/item S_0 to point/item $S_{кон}$). In other tasks these states can be preset completely accurately, but it is only limited by some conditions, i.e., are shown the region of the initial states $\tilde{S}_0$ and the region of final states $\tilde{S}_{кон}$.

The fact that the initial state of system S_0 enters into region $\tilde{S}_0$, we will write/record with the help of the accepted in mathematics "sign of inclusion/connection" $\in$:

$$S_0 \in \tilde{S}_0;$$

it is analogous for the final state:

$$S_{кон} \in \tilde{S}_{кон}.$$

Taking into account initial and final conditions the task of optimum control is formulated as follows: from many possible controls U to find such control U*, which translates the physical system S from initial state $S_0 \in \hat{S}_0$ into final $S_{\kappa} \in \hat{S}_{\kappa}$ so that certain criterion W(U) would be converted into the maximum.

Let us give to control process geometric interpretation. For this for us it is necessary to somewhat widen our customary geometric representations and to introduce the concept about the so-called "phase space"..:.

Page 41.

The state of the physical system S, which we manage, always can be described with the help of one or the other quantity of numerical parameters. Such parameters can be, for example, the coordinates of physical body and its velocity; a quantity of means, included into the group of enterprises; the number of grouping of the troops, etc. These parameters we will call the phase coordinates of system, and the state of system represent in the form of point S' with these coordinates in certain conditional phase space. A change of the state of system S during the control process will be represented as the displacement of point S in the phase space. The selection of control U indicates the selection of the specific trajectory of point S in the phase space, the specific law of motion.

Phase space can be different depending on a number of parameters, which characterize the state of system.

Let, for example, the state of system S be characterized only by the one parameter - coordinate x. Then a change in this coordinate will be represented as the displacement of point S along the axis Ox (or on its specific section, if to coordinate x they are superimposed are some limitations). In this case phase space will be one-dimensional and is the axis of abscissas Ox or its section, and control is interpreted by the law of the motion of point S from initial state $S_0 \in \tilde{S}_0$ into final $S_{кон} \in \tilde{S}_{кон}$ (Fig. 6.1).

If the state of system S is characterized by two parameters x_1 and x_2 (for example, the abscissa of material point and its velocity), then phase space will be plane $x_1 0 x_2$ or its some part (if to parameters x_1 and x_2 are superimposed limitations), and the controlled process will be represented as the displacement of point S from $S_0 \in \tilde{S}_0$ in $S_{кон} \in \tilde{S}_{кон}$ over the specific trajectory on plane $x_1 0 x_2$ (Fig. 6.2).

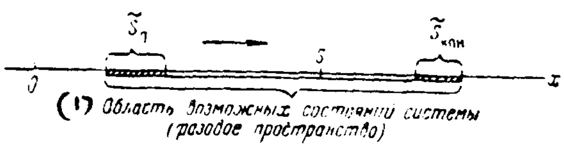


Fig. 6.1.

Key: (1). Region of the possible states of system (phase space).

Page 42.

If the state of system is characterized by three parameters x_1, x_2, x_3 (for example, two coordinates and velocity), then phase space will be ordinary three-dimensional space or its part, and the controlled process will be depicted as the displacement of point S over space curve (Fig. 6.3).

If a number of parameters, which characterize the state of system, is more than three, then geometric interpretation loses its graphic nature, but geometric terminology continues to remain convenient. In general, when the state of system S is described by n by the parameters:

$$x_1, x_2, \ldots, x_n,$$

we will represent it as point S in the n-dimensional phase space, and control interpret as the displacement of point S from some initial region S_0 into final $\tilde{S}_{[illegible]}$ over certain "trajectory", over the specific law.

In order to make clearer an idea of "phase space", let us return to the already examined specific problems which we solved in the previous paragraphs, and let us construct for each of them phase space.

In the task of the optimum gain of altitude and velocity (§3) the state of the physical system S (flight vehicle) was characterized by two phase coordinates - velocity V and with a height of H.

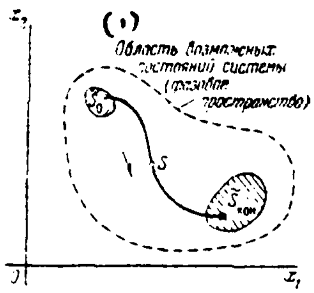


Fig. 6.2.

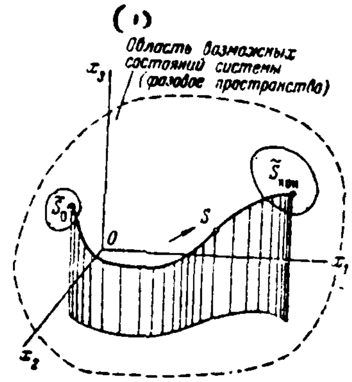


Fig. 6.3.

Fig. 6.2.

Key: (1). Region of the possible states of system (phase space).

Fig. 6.3.

Key: (1). Region of the possible states of system (phase space).

Page 43.

Respectively phase space was two-dimensional and was the phase plane VOH (more accurate, the rectangle, limited by abscissas V_0, $V_{кон}$ and ordinates H_0, $H_{кон}$). Optimum control was represented as the displacement of point S over the optimum trajectory on the phase plane. The initial state $S_0(V_0, H_0)$ and final state $S_{кон}(V_{кон}, H_{кон})$. they were completely determined and were two points S_0, $S_{кон}$ on the phase plane.

In the task of the selection of the fastest path (§4) the physical system S was the truck which was to be translated from initial point S_0 into final $S_{кон}$ within the short time. Here again the state of system was described by two parameters x and y (ordinary Cartesian plane coordinates), and the trajectory of point S on the phase plane was itself the ordinary trajectory of the moving/driving point (truck).

The continuous task of plotting of optimum route, examined in §5, on the setting and in the dimension of space differs in no way from previous. Let us note that if the optimum path ran itself not in the plane, but in the space (for example, the path of aircraft of one point/item in another), then phase space would become three-dimensional, but if in this case was optimized even the mode/conditions of a change in the velocity - four-dimensional.

In all, until now, examples examined initial and final conditions S_0 and $S_{кон}$ were the completely fixed points of phase space. It is not difficult to give examples of the tasks, where these states are the whole regions S_0 and $\tilde{S}_{кон}$ of phase space.

Assume, for example, we need to direct the combat guided rocket from some point S_0 on the earth's surface into the vicinity of target Ts so that it would strike this target. It is obvious, for this there is no necessity to direct rocket into the fixed point S_u and it is sufficient so that it would hit the preset zone $\tilde{S}_u$ which surrounds the target, sizes/dimensions and form of which are determined by damaging effect of rocket. The state of rocket at each moment of time we will as representative point S in the six-dimensional phase space (three coordinates, three components of velocity). At the initial moment the coordinates of rocket are preset; the velocity components are equal to zero (point S_0 is completely determined).

Page 44.

As far as final state is concerned $S_{кон}$, than it is determined not completely: space coordinates must be within the limits of the preset zone $\tilde{S}_u$ and to the components of velocity no limitations are imposed. Consequently, region $\tilde{S}_{кон}$ in the six-dimensional phase space is limited on coordinates x, y, z and is unconfined on coordinates V_x, V_y, V_z.

Let us assume now that the discussion deals not with combat, but about the passenger rocket; for it the touchdown point is completely determined, but on the velocity components are superimposed the severe limitations; in this case region $S_{кон}$ substantially becomes narrow.

In the following presentation we will meet with a whole series of the practical tasks, where the initial state $\tilde{S}_0$ and final $\tilde{S}_{кон}$ are not points, but the whole regions of phase space.

Thus, let us formulate general problem of optimum control in the terms of phase space.

To find such control U* (optimum control), under the effect of which point S of phase space will move from the initial region $\tilde{S}_0$ to finite domain $\tilde{S}_{кон}$ so that in this case criterion W will become maximum.

Stated general/common/total problem can be solved by different methods - far not only by the method of dynamic programming. Characteristic for the dynamic programming is the specific systematic method, namely: the process of the displacement of point S from $\tilde{S}_0$ in $\tilde{S}_{кон}$ is divided into the series/row of consecutive stages (steps/pitches) (Fig. 6.4), and is produced the consecutive optimization of each of them, beginning from the latter.

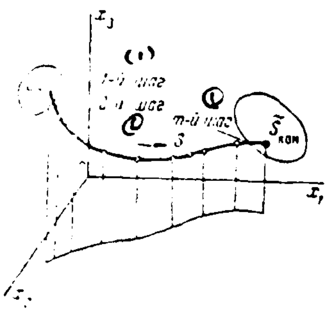


Fig. 6.4.

Key: (1). step/pitch.

Page 45.

In each stage of calculation is sought first conditional optimum control (under all possible assumptions about the results of the previous step/pitch), and then, after the process of optimization is led to the initial state S_0, again passes entire/all sequence of steps/pitches, on already from the beginning to the end/lead, and at each step/pitch from many conditional optimum controls is chosen one.

However, what do we gain with the help of this step-by-step calculation of the process of optimization? We win that the fact that at one step/pitch the structure of control, as a rule, proves to be more simply than for entire elongation/extent of process. Instead of one time solving of complex problem, we prefer many times to solve problem relatively simple.

In this - entire entity of the method of dynamic programming and entire justification for its use/application in practice. If this simplification in the procedure of optimization from the distribution of process into the stages it does not occur, the use/application of a method of dynamic programming becomes meaningless.

§7. General/common/total formula recording of the solution of the problem of optimum control by the method of dynamic programming.

In the previous paragraph we formulated the general/common/total formulation of the problem of dynamic programming and it gave to this task geometric interpretation, after posing it as the problem of steering of point in the phase space.

In the present paragraph we will attempt to register in general form not only setting, but also solution of the problem of dynamic programming. Are true, the formulas, which we will obtain, they will by necessity take the very general/common/total, unspecific form, but for the understanding of the future these general formulas will prove to be useful.

Before to begin the general/common/total formula recording of the process of dynamic programming, we should make more precise the nature of criterion W, with which we thus far in no way dealt.

Page 46.

Let us note that in all examined, until now, examples criterion W possessed one remarkable property: the value of this criterion, achieved/reached for entire process, was obtained by the simple addition of the particular values of the same criterion w_i of achievements at the single steps/pitches.

Actually/really, general/common/total fuel consumption R to the gain of altitude and velocity (see §3) was the sum of fuel consumption r_i at the single steps/pitches:

$$R = \sum_{i=1}^{m} r_i; \qquad (7.1)$$

the total time T of displacement of one point/item in another (see §4) was the sum of the times of overcoming single stages t_i:

$$T = \sum_{i=1}^{m} t_i \qquad (7.2)$$

and so on.

If criterion W possesses this property:

$$W = \sum_{i=1}^{m} w_i, \qquad (7.3)$$

i.e. it is composed of the elementary values of the same criterion, obtained at the single steps/pitches, then it is called additive.

In the majority of the practical tasks, solved by the method of dynamic programming, criterion was additive. If it in the initial formulation of the problem is not additive, then they try then to modify this setting or criterion itself so that it would gain the property of the additivity (see further, §14).

We will examine only the additive tasks of dynamic programming and some most elementary ones from those leading to additive.

Let us give setting and overall diagram of the solution of the problems of dynamic programming with the additive criterion.

Let there be the control process of the physical system S, separated on m of steps/pitches (stages).

Page 47.

At our disposal at each (i-th) step/pitch is control U_i by means of which we we translate system of the state S_{i-1} of the achievement as a result (i-1) step/pitch, into new state S_i which depends on S_{i-1} and selected by us control U_i. This dependence we will register as follows:

$$S_i = S_i(S_{i-1}, U_i). \qquad (7.4)$$

by considering S_i as the function of two arguments S_{i-1} and U_i

Let us note that for applying the method of dynamic programming it is substantial so that the new state S_i would depend only on state S_{i-1} and control at i step/pitch U_i and it did not depend on how system arrived into state S_{i-1}. If this proves to be not then, then should be "enriched" the concept of the "state of system", after introducing into it those parameters from the past, on which depends the future, i.e., to increase a number of measurements of phase space.

Under the effect of controls $U_1, U_2, \ldots, U_m$ the system passes from the initial state S_0 into final $S_{кон}$. As a result of entire process after m of steps/pitches is obtained the "income" or "prize"

$$W = \sum_{i=1}^{m} w_i(S_{i-1}, U_i). \qquad (7.5)$$

where $w_i(S_{i-1}, U_i)$ - prize at the i step/pitch, which depends, naturally, on the previous state of system S_{i-1} and selected control U_i.

Is preset the region of the initial states $\tilde{S}_0$ and the final states $\tilde{S}_{кон}$. It is necessary to select initial state $S_0^* \in \tilde{S}_0$ and controls $U_1^*, U_2^*, \ldots, U_m^*$ at each step/pitch so that after m steps/pitches system would pass into region $\tilde{S}_{кон}$ and in this case prize W became maximum.

Let us describe in general form the procedure of the use/application of a method of dynamic programming to the solution of this problem.

For this by us it will be necessary to introduce some new designations. We already designated W - prize within always of process; w_i - prize for the i step/pitch. Since the process of dynamic programming is turned/run up from the end/lead, for us it is necessary to introduce special designation for the prize, acquired for several latter/last steps/pitches of process.

Page 48.

Let us designate:

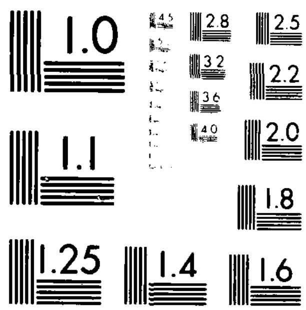
W_m - prize for the latter/last step/pitch,

$W_{m-1,m}$ - prize for two latter/last steps/pitches,

$W_{i,i+1,\ldots,m}$ - prize for the latter (m-i+1) of step/pitch, beginning from the i-th and ending with the m-th.

AD-A098 406 FOREIGN TECHNOLOGY DIV WRIGHT-PATTERSON AFB OH F/G 12/1
ELEMENTS OF DYNAMIC PROGRAMMING.(U)
FEB 81 Y S VENTTSEL'
UNCLASSIFIED FTD-ID(RS)T-1515-80 NL

2 of 4
AD
A098406



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

It is obvious,

$$\left.\begin{array}{l} W_m = w_m, \\ W_{m-1,m} = w_{m-1} + w_m, \\ \dots\dots\dots\dots\dots\dots\dots \\ W_{i,i+1,\dots,m} = w_i + w_{i+1} + \dots + w_m. \end{array}\right\} \quad (7.6)$$

As we already know, the process of the optimization of control of the method of dynamic programming begins from the last (m-th) step/pitch. Let afterward (m-1)-th step/pitch the system be in state S_{m-1}. Since latter/last (m-th) step/pitch must translate system into state $S_m = S_{кон} \in \tilde{S}_{кон}$, then as S_{m-1} it is possible to take not all in the principle possible states of system, but only those from which for one step/pitch it is possible to pass into region $\tilde{S}_{кон}$.

Let us assume that state S_{m-1} to us is known, and let us find under this condition conditional optimum control on the m step/pitch; let us designate it $U_m^*(S_{m-1})$. This is - the control which, being used at the m step/pitch, translates system into final state $S_m \in \tilde{S}_{кон}$, the prize at this latter/last step/pitch W_m reaching its maximum value:

$$W_m^*(S_{m-1}) = \max_{U_m} \{W_m(S_{m-1}, U_m)\}. \quad (7.7)$$

Let us recall the sense of symbolic formula (7.7). $W_m(S_{m-1}, U_m)$ indicates prize generally (not optimum) at the latter/last step/pitch; it depends both on the result of previous step/pitch S_{m-1} and on used at this step/pitch control U_m. Of all gains

$W_m(S_{m-1}, U_m)$ during different controls U_m is chosen that prize $W^*(S_{m-1})$, which has maximum value; this it indicates recording $\max_{U_m}$.

Page 49.

Let us note that as controls U_m we must take only those which translate system of the present state S_{m-1} into state S_m, belonging to region $\tilde{S}_{кон}$.

Finding the conditional maximum value of prize $W^*_m(S_{m-1})$, we thereby find conditional optimum control $U^*_m(S_{m-1})$. The fact that conditional maximum prize $W^*_m(S_{m-1})$ is achieved by conditional optimum control $U^*_m(S_{m-1})$, we will register symbolically in the form

$$W^*_m(S_{m-1}) \sim U^*_m(S_{m-1})$$

and subsequently we will use this recording for the indication of conformity between the conditional maximum prize and the conditional optimum control at each step/pitch.

Thus, the optimization of latter/last step/pitch with any result of next-to-last is produced, and is found the corresponding conditional optimum control. The obtained result can be formulated thus: in whatever state proved to be the system after (m-1) step/pitch, we already know that to us to make further.

Let us switch over to the optimization of next-to-last (m-1) step/pitch. Let us do again the assumption that as a result (m-2) step/pitch the system arrived into state S_{m-2}. Let at (m-1) step/pitch we use control U_{m-1}. As a result of this control we at (m-1) step/pitch will obtain the prize, which depends both on the state of system and on the used control:

$$w_{m-1} = w_{m-1}(S_{m-2}, U_{m-1}). \qquad (7.8)$$

and system becomes new state S_{m-1}, also depending on the previous state and on the control:

$$S_{m-1} = S_{m-1}(S_{m-2}, U_{m-1}). \qquad (7.9)$$

But for any result of (m-1) step/pitch the following, m step/pitch is already optimized, and maximum prize on it is equal to

$$W_m^*(S_{m-1}) = W_m^*(S_{m-1}(S_{m-2}, U_{m-1}))^1. \qquad (7.10)$$

FOOTNOTE [1]. The sense of recording (7.10) following: prize W_m^* there is a function of state S_{m-1}, which in turn, depends on previous state S_{m-2} and used control U_{m-1}. Since for the designation of functional dependence it is accepted to use the parenthesis, then in the formulas of type (7.10) we place the parenthesis of inside circular ones. ENDFOOTNOTE.

Page 50.

Let us introduce into the examination full/total/complete prize at two latter/last steps/pitches during any control at (m-1) step/pitch optimum control at the m step/pitch. Let us designate its $W^+_{m-1,m}$; sign "+" it will us remind that this is prize during the incompletely optimized control, in contrast to the sign "*", by which we designated prize with that completely optimized control. Prize $W^+_{m-1,m}$ it is obvious, it depends on the previous state of system S_{m-2} and used at (m-1) step/pitch control U_{m-1}. Taking into account formulas (7.8) and (7.10), we will obtain the following expression for $W^+_{m-1,m}$:

$$W^+_{m-1,m}(S_{m-2}, U_{m-1}) = \\ = w_{m-1}(S_{m-2}, U_{m-1}) + W^*_m(S_{m-1}(S_{m-2}, U_{m-1})). \quad (7.11)$$

We should select this optimum conditional control at (m-1) step/pitch $U^*_{m-1}(S_{m-2})$, with which value (7.11) would achieve the maximum:

$$W^*_{m-1,m}(S_{m-2}) = \max_{U_{m-1}} \{W^+_{m-1,m}(S_{m-2}, U_{m-1})\}. \quad (7.12)$$

Just as in the previous stage of optimization, as states S_{m-2} afterward (m-2) steps/pitches it is necessary to take not all possible states of system, but only those from which it is possible to pass in $\tilde{S}_{кон}$ for two steps/pitches.

Thus, is found maximum conditional prize on two latter/last steps/pitches and corresponding to it optimum conditional control at (m-1) step/pitch:

$$W^*_{m-1,m}(S_{m-2}) \sim U^*_{m-1}(S_{m-2}).$$

By continuing by accuracy such form, it is possible to find conditional maximum prizes on several latter/last steps/pitches of process and corresponding to them optimum conditional controls:

$$W^{*}_{m-2,\, m-1,\, m}(S_{m-3}) \sim U^{*}_{m-2}(S_{m-3}),$$
$$W^{*}_{m-3,\, m-2,\, m-1,\, m}(S_{m-4}) \sim U^{*}_{m-3}(S_{m-4})$$

and so forth.

Page 51.

If we already optimized (i+1) step/pitch for any issue of the i-th, i.e., they found

$$W^{*}_{i+1,\, \ldots,\, m}(S_{i}) \sim U^{*}_{i+1}(S_{i}),$$

that the conditional optimization of the i step/pitch it is produced according to the general formula

$$W^{*}_{i,\, i+1,\, \ldots,\, m}(S_{i-1}) = \max_{U_i} \{W^{+}_{i,\, i+1,\, \ldots,\, m}(S_{i-1},\, U_i)\}, \quad (7.13)$$

where

$$W^{+}_{i,\, i+1,\, \ldots,\, m}(S_{i-1},\, U_i) = \\ = w_i(S_{i-1},\, U_i) + W^{*}_{i+1,\, \ldots,\, m}(S_i(S_{i-1},\, U_i)) \quad (7.14)$$

- prize, reached at the latter/last steps/pitches, beginning from the i-th, during any control at the i step/pitch and optimum control on all those following; $S_i(S_{i-1}, U_i)$ - the state into which passes the system from S_{i-1} under the effect of control U_i.

Thus, is determined conditional maximum prize at the last steps/pitches, beginning from the i-th, and the corresponding optimum conditional control at the i step/pitch:

$$W^*_{i,\, i+1,\, \dots,\, m}(S_{i-1}) \sim U^*_i(S_{i-1}). \qquad (7.15)$$

Applying consecutively/serially, step by step, the described procedure, we will reach finally the first step/pitch:

$$W^*_{1,\, \dots,\, m}(S_0) \sim U^*_1(S_0). \qquad (7.16)$$

where S_0 - some initial state of system, which belongs to region $\tilde{S}_0$ of the possible initial states: $S_0 \in \tilde{S}_0$.

Remains to select optimally the initial state of system S^*_0. if the initial state S_0 in the accuracy preset (i.e. entire/all region $\tilde{S}_0$ is reduced to one point S_0), then there is no selection, and $S^*_0 = S_0$. But if point S_0 can freely be chosen in the limits of region $\tilde{S}_0$, then it is necessary to optimize the selection of initial state, i.e., to find absolute (no longer conditional) maximum prize in all steps/pitches:

$$W^*_{1,\, \dots,\, m} = \max_{S_0 \in \tilde{S}_0} \{W^*_{1,\, \dots,\, m}(S_0)\}. \qquad (7.17)$$

where the recording $\max\limits_{S_0 \in \tilde{S}_0}$ indicates: maximum is taken due to all states S_0, entering region $\tilde{S}_0$. Point S^*_0, at which it is reached this maximum, and should be taken as the initial state of system.

Page 52.

Thus, as a result of the consecutive passage of all stages from the end/lead at the beginning are found: the maximum value of prize for all m steps/pitches and corresponding to it optimum initial state of the process

$$W^* = W^*_{1,2,\ldots,m} \sim S^*_0. \qquad (7.18)$$

But is constructed already optimum control? No yet: indeed we found on each step/pitch only an conditional optimum control.

In order to find optimum control in the final instance, we must again pass entire sequence of steps/pitches - this time from the beginning toward the end. This second "passage on the steps/pitches" will be much simpler than the first, because to vary conditions no longer it is necessary.

As the initial state of system is selected S^*_0 (or simply S_0, if initial state is rigidly fixed/recorded). At the first step/pitch is applied the optimum control u^*_1 (see (7.16))

$$U^*_1 = U^*_1(S^*_0). \qquad (7.19)$$

after which system it passes into newly the state

$$S^*_1 = S_1(S^*_0, U^*_1). \qquad (7.20)$$

It is now necessary to select optimum control at the second step/pitch. We already optimized it for any result of the first step/pitch, i.e. we knew $U^*_2(S_1)$ (see (7.15)); substituting in it S^*_1, we will obtain

$$U^*_2 = U^*_2(S^*_1). \qquad (7.21)$$

and so on, until we reach the optimum control at the latter/last step/pitch

$$U^*_m = U^*_m(S^*_{m-1}) \qquad (7.22)$$

and the final states of the system

$$S^*_m = S^*_{\text{кон}} = S^*_m(S^*_{m-1}, U^*_m). \qquad (7.23)$$

As a result of this entire procedure is found finally the solution of the problem: maximally possible prize W* and the optimum control U* which consists of the optimum controls on the single steps/pitches (vector of optimum control)

$$U^* = (U^*_1, U^*_2, \ldots, U^*_m). \qquad (7.24)$$

Page 53.

Thus, in the process of dynamic programming the sequence of stages passes twice: for the first time - from the end/lead at the beginning, as a result of which is found the maximum value of prize W*, the optimum initial state of process S^*_0 and conditional optimum control at each step/pitch; for the second time - from the beginning toward the end, as a result of which is found optimum control U^*_i on each step/pitch and final state of system with the optimum control $S^*_{\text{кон}}$.

Thus, we succeeded in presenting in general form and registering with the help of the general formulas the process of dynamic programming. In view of the symbolic recording of formulas the structure of process is very simple, but this - false impression. Upon the setting of the specific problems of dynamic programming frequently appear the difficulties.

These difficulties are connected, in the first place, with the selection of the group of parameters x_1, x_2, ..., x_n, characterizing the state of the physical system S. As it was already said, these parameters must be chosen so that in the preset state $S(x_1, x_2, \ldots, x_n)$ of system S at any moment/torque its following state $S'(x'_1, x'_2, \ldots, x'_n)$, into which it passes under the effect of control U, it depended only on the previous state S and control U and did not depend on "past history" of process, i.e., from that, when, as as a result of what controls system arrived into state S. If this proves to be not then it comes those elements/cells of the past on which depends the future, to include in the set of parameters $x_1, x_2, \ldots, x_n$, those characterizing the state of system at the given moment/torque. But this leads to an increase in the number of measurements of phase space and, which means, to the complication of task.

The second difficulty consists of reasonable "staging" of process. It is necessary so to disengage the process of transition from $\tilde{S}_0$ to $\tilde{S}_{\text{кон}}$ to the steps/pitches so that they would allow/assume the convenient numbering and the precise sequence of operations. This task is frequently far not simple.

Page 34.

As has been mentioned above, the distribution of process into the discrete/digital "steps/pitches" is not the necessary sign/criterion of the method of dynamic programming. In principle always it is possible to direct the length of step/pitch toward zero and to consider limiting case - "continuous" dynamic programming. It is possible to obtain in the final form the solution of such continuous problems only in the rare cases, but they have the high theoretical value in the proof of the existence theorems, and also different qualitative properties of the optimum solutions (see [1]). In our elementary presentation of the method of dynamic programming we will in no way concern these limiting cases.

In further paragraphs we will consider a whole series of the practical tasks, which are adequate/approach under the overall diagram of dynamic programming. Some of these tasks we will only supply, for the majority let us sketch the diagram of the solution, while some solution to the end/lead. Some tasks comparatively easily are carried out under the overall diagram, presented in this paragraph; above the setting of others it is necessary to still take some pains itself. Keeping in mind unwieldiness of calculations "by hand", it is easy to comprehend that to the concrete/specific/actual numerical result will be led only the simplest tasks with a small number of parameters x_1, x_2, ..., that are determining the state of system. However, it is necessary to have in mind that by completely the same methods with the help of the contemporary high speed computers it is possible to solve and much more complex problems with a considerable number of parameters. However, as far as number is concerned of steps/pitches m, then for the machine calculation its increase difficulties does not generally obtain: simply increases the time of calculation proportional to a number of steps/pitches.

Page 55.

§3. Task of distributing the resources/lifetimes.

Among the practical tasks, solved by the method of dynamic programming, many they have as a goal to find the rational distribution of resources/lifetimes according to different categories of actions. To this type belongs, for example, the task about the distribution of means to the equipment, the purchase of raw material and the hire of work force during the organization for work of industrial enterprise; the task about the distribution of goods according to the commercial and storage locations; the task about the distribution of means between different branches of industry; the task about the weight distribution between different aggregates/units of technical device/equipment, etc.

Here we will consider one of the simplest tasks of distributing the resources/lifetimes, on which it is easy to demonstrate the special feature/peculiarity of many similar tasks.

There is preset initial quantity of means Z_0 (it is not compulsory in the money form), which we must distribute between two branches of the production: I and II. These means, being they are imbedded in branch I and II, yield the specific income. A quantity of means x, included into branch I, in one year arrives income f(x); in this case it is reduced (partially it is expended), so that toward the end of the year from it remains the remainder/residue, equal to $\phi(x)$:

$$\varphi(x) \leqslant x.$$

Is analogous a quantity of means y, imbedded in branch II, yields in the year income g(y) and is reduced to $\psi(y)$:

$$\psi(y) \leqslant y.$$

After each year the remaining means again are distributed between the branches. New means it does not act, and into the production are packed all remaining in the presence means.

It is necessary to find such method of control of service lives

- what means, in what years and into what branches to pack, with which total income during the period into m of years is converted into the maximum.

We will solve problem by the method of dynamic programming. The physical system S, which we will manage, is the group of enterprises with the imbedded in them means. Prize W - income from both branches I and II during entire period. Assignment - to plan/glide on m of years - gives the natural articulation of process on m of steps/pitches (stages). However, for the purpose of presentation we will at each step/pitch distinguish two halfsteps or "component/link". On the first of them occurs the redistribution of means; on the second - the means only are expended and occurs formation of income.

Page 56.

Let us select the now numerical parameters with the help of which we will characterize situation (state of system).

The situation before beginning the i stage (before the redistribution of means) let us agree to characterize by quantities of the means

$$x'_{i-1}, \quad y'_{i-1}.$$

those remaining in branches I and II after previous (i-1) step/pitch[1].

FOOTNOTE [1]. This does not relate to the first step/pitch in the beginning of which to us is simply given certain quantity of means Z_0. ENDFOOTNOTE.

Situation after the distribution of means (i.e. after the first component/link of the i step/pitch) we will characterize by quantities of the means

$$x_i, \quad y_i,$$

those packed in branch I and II at this step/pitch.

As a result of the second component/link of the i step/pitch (consumption of resources) these values it is reduced and will become equal to

$$x_i', \; y_i',$$

after which we let us pass to the following step/pitch.

Let us depict the state of system as point S in the phase space. This space can be constructed in different ways; for the purpose of clarity we will select its two-dimensional (Fig. 8.1).

Along the axis of abscissas ox we will plot/deposit the quantity of resources, packed into branch I; on axis of ordinates Oy - quantity of resources, packed into branch II. Then phase space will be the part of plane xOy, which lies within and on the borders of triangle AOB. Actually/really, for any stage of production the sum of the resources, imbedded in branch I and II, cannot exceed the initial supply of the resources:

$$x + y \leqslant Z_0; \qquad (8.1)$$

furthermore, these enclosures are non-negative:

$$x \geqslant 0; \quad y \geqslant 0. \qquad (8.2)$$

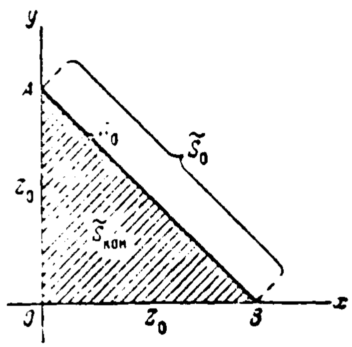


Fig. 3.1.

Page 57.

By the region of plane xOy, which satisfies conditions (8.1) and (8.2), is triangle ACB; this and there is the phase space, in which it can change its position point S, which represents the state of system.

Let us determine regions $\tilde{S}_0$ and $\tilde{S}_{кон}$ of the initial and final conditions of system.

At the initial moment single, that we know about the state of system, this that the fact that the sum of enclosures into both branches is equal to the initial supply of the resources:

$$x+y=Z_0.$$

This condition satisfies any point cutting off AB, which and is region $\tilde{S}_0$ of the initial states of system. As far as position is concerned of end point $\tilde{S}_{\kappa o \mu}$. then we know only that for it

$$x \geqslant 0, \quad y \geqslant 0, \quad x+y < Z_0,$$

i.e. region $\tilde{S}_{\kappa o \mu}$ is entire triangle AOB, besides hypotenuse AB.

Let us look, what form can take the trajectory of point S in the phase space. Since we examine discretized problem, this trajectory we will represent in the form of broken line (Fig. 8.2). At the first step/pitch, in contrast to others, occurs only the expenditure of the resources (there is no redistribution). In this case of the point S_0 with the coordinates (x_1, y_1) we pass into point M with the coordinates (x'_1, y'_1). Since $x'_1 \leqslant x_1$, $y'_1 \leqslant y_1$, the this component/link of trajectory is the segment, directed from point S_0 downward and to the left. The following (second) step/pitch is divided into two components/links: 2_1 and 2_2. On the first component/link 2_1 occurs the redistribution of resources; in this case x+y remains constant and, which means, point S is moved on the straight line, parallel AB, into point N with the coordinates (x_2, y_2). On the second component/link of second step/pitch (2_2) again occurs the expenditure of resources, and point S moves downward and to the left, and so on, until through m of steps/pitches is achieved/reached final state $S_{\kappa o \mu}$ - point with coordinates (x'_m, y'_m).

Page 58.

Let us note that the components/links of steps/pitches are nonequivalent: control is realized only on the first component/link of each step/pitch, and on the second we obtain income. Control U_i on the i step/pitch (realized on the first component/link i_1) consists of the selection of non-negative values x_i, y_i of such, that

$$x_i + y_i = x'_{i-1} + y'_{i-1}.$$

After this we obtain on the second component/link of the i step/pitch (I_2) the income

$$w_i = f(x_i) + g(y_i). \qquad (8.3)$$

but point S, which represents the state of system, passes to the new position with the coordinates

$$x'_i = \varphi(x_i); \quad y'_i = \psi(y_i). \qquad (8.4)$$

It is necessary to find this position S^*_0 of point S_0 on the straight line AB and this trajectory of point S in the phase space so that the total income for all m of the steps/pitches

$$W = \sum_{i=1}^{m} w_i \qquad (8.5)$$

would be converted into the maximum.

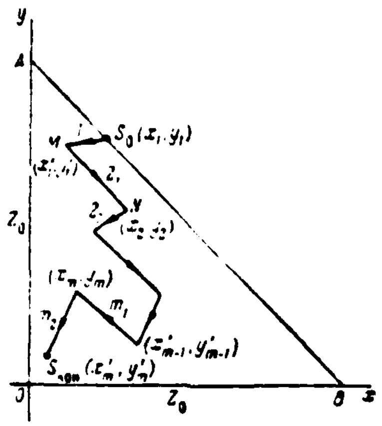


Fig. 8.2.

Page 59.

Before us - the typical task of dynamic programming. let us use to its solution the general/common/total methods, presented in the previous paragraph. In order to make a concrete/specific/actual application/appendix of general method as clear as possible, we will permit ourselves perhaps a little to be repeated.

As always, we will optimize the process of distributing the resources, beginning from the end/lead, moreover immediately on both components/links of each step/pitch (taking into account that the second of them it is unguided).

Let before the m-th (latter/last) step/pitch we be found at point (x'_{m-1}, y'_{m-1}), and we must redistribute resources, i.e., select point (x_m, y_m) such, that

$$x_m + y_m = x'_{m-1} + y'_{m-1}.$$

Let us note that for solving this for us is not required knowledge of both numbers x'_{m-1}, y'_{m-1}, and it is essential to know only their sum, which is subject to the redistribution:

$$Z_{m-1} = x'_{m-1} + y'_{m-1}\ [1].$$

FOOTNOTE [1]. Since the state of system after each stage is characterized only by one number, we could select our phase space one-dimensional, but then trajectory would appear so not clearly. ENDFOOTNOTE.

Redistribution will consist in the fact that we will isolate some part x_m of resources Z_{m-1} and put it into branch I; a quantity of resources y_m, which is packed into branch II, automatically it will be determined from the relationship/ratio

$$y_m = Z_{m-1} - x_m.$$

Thus, at m step/pitch "control" is x_m. We must find on this step/pitch conditional optimum control, i.e., for any value Z_{m-1} find such quantity of resources $x^*_m(Z_{m-1})$, of those packed in branch I, with which the income at the m step/pitch, equal to

$$W_m(Z_{m-1}, x_m) = w_m(Z_{m-1}, x_m). \qquad (8.6)$$

is converted into the maximum:

$$W^*_m(Z_{m-1}) = \max_{0 < x_m < Z_{m-1}} \{W_m(Z_{m-1}, x_m)\}. \quad (8.7)$$

Page 60.

Recording max. means that is taken the maximum

$$0 < x_m \leq Z_{m-1}$$

in terms of all possible at this step values of control x_m; they are non-negative and do not exceed the general/common/total supply of means Z_{m-1}, by which we arrived at this step.

Expressing maximum income (8.7) at the last step/pitch through the imbedded means according to formula (8.3), we will obtain

$$W^*_m(Z_{m-1}) = \max_{0 \leq x_m \leq Z_{m-1}} \{f(x_m) + g(Z_{m-1} - x_m)\}. \quad (8.8)$$

To this maximum value corresponds the specific conditional optimum control at the m step/pitch

$$W^*_m(Z_{m-1}) \sim U^*_m(Z_{m-1}).$$

and the problem of the conditional optimization of the m step/pitch is solved.

Let us switch over to the conditional optimization of next-to-last ((m-1)-th step/pitch. Let after (m-2)-th step/pitch be preserved the supply of the means

$$Z_{m-2} = x'_{m-2} + y'_{m-2}. \quad (8.9)$$

Let us find $W^{*}_{m-1,m}(Z_{m-2})$ - conditional maximum income in two latter/last steps/pitches. Let control U_{m-1} used at (m-1)-th step/pitch, consist of the fact that we pack into branch I provisions x_{m-1} (and that means, into branch II - provisions $y_{m-1}=Z_{m-2}-x_{m-1}$). With respect to these enclosures at (m-1)-th step/pitch we will obtain the income

$$w_{m-1}(Z_{m-2}, x_{m-1})=f(x_{m-1})+g(Z_{m-2}-x_{m-1}). \quad (8.10)$$

and system changes into the point of phase space with the coordinates

$$x'_{m-1}=\varphi(x_{m-1});\ y'_{m-1}=\psi(y_{m-1})=\psi(Z_{m-2}-x_{m-1}). \quad (8.11)$$

According to the general/common/total principle (see §7) in order to optimize conditional control at (m-1)-th step/pitch, it is necessary to sum income at (m-1)-th step/pitch (8.10) during any control x_{m-1} with the already optimized income at m step/pitch (8.8); we will obtain total income at two latter/last steps/pitches

$$W^{+}_{m-1,m}(Z_{m-2}, x_{m-1})=$$
$$=w_{m-1}(Z_{m-2}, x_{m-1})+W^{*}_{m}(Z_{m-1}). \quad (8.12)$$

Page 61.

After this let us find the control x_{m-1} on (m-1)-th step/pitch with which income (8.12) is converted into the maximum:

$$W^{*}_{m-1,m}(Z_{m-2})=$$
$$=\max_{0\leq x_{m-1}\leq Z_{m-2}}\{W^{+}_{m-1,m}(Z_{m-2}, x_{m-1})\}. \quad (8.13)$$

Let us write evident expression $W^{+}_{m-1,m}(Z_{m-2}, x_{m-1})$ as to function from both arguments. For this let us substitute into formula (8.12) expression (8.10):

$$W_{m-1,m}(Z_{m-2}, x_{m-1}) = = f(x_{m-1}) + g(Z_{m-2}, -x_{m-1}) + W^{*}_{m}(Z_{m-1}). \quad (8.14)$$

But on right side of (8.14) is included, besides Z_{m-2} and x_{m-1}, still Z_{m-1}. In order to get rid of this "excess" argument, let us recall that the supply of means Z_{m-1} after (m-1)-th step/pitch depends on the supply of means Z_{m-2}, of available at the beginning of this step/pitch, and used at (m-1)-th step/pitch control x_{m-1}, according to formula (8.4)

$$Z_{m-1} = \varphi(x_{m-1}) + \psi(Z_{m-2} - x_{m-1}). \quad (8.15)$$

Substituting this expression into formula (8.14) and then (8.14) in (8.13), we will obtain finally the expression of conditional maximum income at two latter/last steps/pitches:

$$W^{*}_{m-1,m}(Z_{m-2}) = = \max_{0 \leq x_{m-1} \leq Z_{m-2}} \{f(x_{m-1}) + g(Z_{m-2} - x_{m-1}) + + W^{*}_{m}(\varphi(x_{m-1}) + \psi(Z_{m-2} - x_{m-1}))\}. \quad (8.16)$$

where f, g, φ ψ - specific, preset functions of their arguments, and $W^{*}_{m}(Z_{m-1})$ - function, obtained as a result of the conditional optimization of latter/last step/pitch; into this function (given cne by formula, graph or table) instead of argument Z_{m-1} it is necessary to substitute value (8.15).

Page 62.

The value x_{m-1}, with which attains maximum (8.16), and there is conditional optimum control at (m-1)-th step/pitch

$$x^*_{m-1}(Z_{m-2}).$$

Thus, the problem of the conditional optimization of control at (m-1)-th step/pitch is solved: is found conditional maximum income in two latter/last steps/pitches and corresponding to it conditional optimum controls - quantity of means, packed on (m-1)-th step/pitch into branch I:

$$W^*_{m-1,m}(Z_{m-2}) \sim x^*_{m-1}(Z_{m-2}).$$

By continuing the process of conditional optimization in exactly the same manner, we will obtain for any (the i-th) step/pitch conditional maximum income for all steps/pitches, by beginning from the data:

$$W^*_{i,i+1,\ldots,m}(Z_{i-1}) = \\ = \max_{0 \leq x_i \leq Z_{i-1}} \{W^+_{i,i+1,\ldots,m}(Z_{i-1}, x_i)\}, \quad (8.17)$$

where

$$W^+_{i,i+1,\ldots,m}(Z_{i-1}, x_i) = f(x_i) + g(Z_{i-1} - x_i) + \\ + W^*_{i+1,\ldots,m}(\varphi(x_i) + \psi(Z_{i-1} - x_i)). \quad (8.18)$$

and $W^*_{i,\ldots,m}(Z_i)$ - function, already constructed during the optimization of the i step/pitch: into this function instead of argument Z_i it is necessary to substitute the expression

$$\varphi(x_i)+\psi(Z_{i-1}-x_i). \qquad (8.19)$$

Substituting (8.18) in (8.17), we will obtain evident expression $W^*_{i,\ldots,m}(Z_{i-1})$ through known functions $f, g, \varphi, \psi, W^*_{i+1,\ldots,m}$:

$$W^*_{i,i+1,\ldots,m}(Z_{i-1})=\max_{0<x_i\leq Z_{i-1}}\{f(x_i)+g(Z_{i-1}-x_i)+ \\ +W^*_{i+1,\ldots,m}(\varphi(x_i)+\psi(Z_{i-1}-x_i))\}. \qquad (8.20)$$

To this conditional maximum income corresponds conditional optimum control at the i step/pitch:

$$W^*_{i,i+1,\ldots,m}(Z_{i-1})\sim x^*_i(Z_{i-1}). \qquad (8.21)$$

When thus we produce the conditional optimization of all steps/pitches, except the first (let us recall that it is qualitatively different from the others, since it consists only of one component/link), to us it remains to optimize control on this first step/pitch and to find the maximum full/total/complete prize at all steps/pitches, which depends it goes without saying on the initial supply of means Z_0:

$$W^*(Z_0)=W^*_{1,2,\ldots,m}(Z_0). \qquad (8.22)$$

Page 63.

Value $W^{*}_{1,2,\ldots,m}(Z_0)$ will be located from the same formula (8.20) as at the remaining steps/pitches:

$$W^{*}_{1,2,\ldots,m}(Z_0)=\max_{0\leq x_1\leq Z_1}\{f(x_1)+g(Z_0-x_1)+ + W^{*}_{2,\ldots,m}(\varphi(x_1)+\psi(Z_0-x_1))\}. \quad (8.23)$$

Entire/all special feature/peculiarity of the first step/pitch lies in the fact that the initial supply of means Z_0 is not varied, but it is assumed to be known. The value of control x^{*}_1, at which reaches maximum (8.23), is no longer conditionally optimum, but simply optimum control at the first step/pitch which it is necessary to use.

This value x^{*}_1 determines the abscissa of point S^{*}_0 on cutting off AB, with which begins optimum trajectory in the phase space.

Knowing the position of this point and again passing all steps/pitches, but already in the opposite direction - from the beginning toward the end, it is possible to construct entire optimum trajectory of point S. Let us trace how will pass this trajectory, on the steps/pitches and their components/links.

In the beginning of the first step/pitch point S is found on cutting off AB and has coordinates

$$x^{*}_1, \quad y^{*}_1=Z_0-x^{*}_1.$$

After the first step/pitch point S is moved into the point with the coordinates

$$(x'_1)^{*}=\varphi(x^{*}_1); \quad (y'_1)^{*}=\psi(y^{*}_1).$$

sum of which is equal to the supply of means after the first step/pitch

$$Z_1^* = (x_1')^* + (y_1')^*.$$

On the first component/link of the second step/pitch occurs the redistribution of means; point S passes into the point with the coordinates

$$x_2^* = x_2^*(Z_1^*), \quad y_2^* = Z_1^* - x_2^*,$$

where $x^*_2(Z^*_1)$ - the conditional optimum control at the second step/pitch, in which instead of Z_1 is set Z^*_1.

Page 64.

On the second component/link of the second step/pitch occurs the expenditure of the means, and point S is moved into the point with the coordinates

$$(x_2')^* = \varphi(x_2^*); \quad (y_2')^* = \psi(y_2^*),$$

sum of which is equal to the remaining toward the end of the second step/pitch supply of the means

$$Z_2^* = (x_2')^* + (y_2')^*$$

and so forth up to the latter/last step/pitch.

Thus find the final solution of the problem: maximum income for all m of steps/pitches W* and corresponding to it optimum control $X^*=X^*(x^*_1, x^*_2, \ldots, x^*_m)$ indicating, what quantity of means in what stage it is necessary to select into branch I (remainder/residue automatically is abstracted/removed to branch II).

After is examined the specific problem of dynamic programming, it is useful again to return to the general/common/total presentation of a question into §7 and to look, what concrete/specific/actual embodiment obtained in this problem the introduced there general/common/total concepts.

The system of these conformity we will register in the form of the table, divided into two parts by vertical feature; to the left of the feature we will write that value, concept or symbol which was applied in general; to the right - corresponding to it analog in our special case.

Page 65.

(1) В общем случае	(2) В нашем частном случае
(3) Физическая система S	(4) Группа предприятий с вложенными в них средствами
(5) m этапов (шагов)	(6) m лет
(7) Аддитивный критерий $W = \sum_{i=1}^{m} w_i$, где w_i — выигрыш на i-м шаге	(8) Общий доход за m лет $W = \sum_{i=1}^{m} w_i$, где w_i — доход от отраслей I и II на i-м шаге
(9) Управление U_i на i-м шаге	(10) Количество средств x_i, вкладываемое в отрасль I
(11) Состояние системы после i-го шага S_i	(12) Количество средств x_i', y_i', оставшихся в отраслях I и II соответственно, существенной для планирования дальнейших шагов является их сумма $Z_i = x_i' + y_i'$
(13) Состояние системы после i-го шага в зависимости от ее состояния после $(i-1)$-го шага и управления на i-м шаге $S_i = S_i(S_{i-1}, U_i)$	$Z_i = \varphi(x_i) + \psi(Z_{i-1} - x_i)$
(14) Выигрыш на i-м шаге в зависимости от исхода $(i-1)$-го шага S_{i-1} и примененного на i-м шаге управления: $w_i(S_{i-1}, U_i)$	$w_i(Z_{i-1}, x_i) = f(x_i) + g(Z_{i-1} - x_i)$
(15) Фазовое пространство	(16) Треугольник AOB (см. рис. 8.1)
(17) Область начальных состояний системы $\tilde{S}_0$	(18) Отрезок AB (см. рис. 8.1)
(19) Область конечных состояний системы $\tilde{S}_{кон}$	(20) Треугольник AOB (за исключением гипотенузы) (см. рис. 8.1)
Оптимальное начальное состояние системы (21) S_0^*	(22) Оптимальное количество средств x_1^*, выделенное в первую отрасль, и определяемое им количество средств $y_1^* = Z_0 - x_1^*$, выделенное во вторую отрасль

(23) Оптимальное управление $U^* = (U_1^*, U_2^*, \ldots, U_m^*)$	(24) Оптимальное количество средств по годам, выделяемое в отрасль I: $X^* = (x_1^*, x_2^*, \ldots, x_m^*)$

Key: (1). In general. (2). In our special case. (3). Physical system S. (4). Group of enterprises with imbedded in them means. (5). m of ethanes (steps/pitches). (6). years. (7). Additive criterion where w_i - prize at i step/pitch. (8). Aggregate profit after m of years
where w_i - income from branches I and II at i step/pitch. (9). Control at i step/pitch. (10) A quantity of means x_i packed into branch I. (11). State of system afterward i-th step/pitch S_i.(12). Quantity of means x_i', y_i' remaining in branches I and II respectively, essential for planning further steps/pitches is their sum. (13). State of system after i step/pitch depending on its state after (i-1)-th step/pitch and control at i step/pitch. (14). Prize at i step/pitch depending on issue of (i-1)-th step/pitch S_{i-1} and used at i step/pitch control. (15). Phase space. (16). Triangle AOB (see Fig. 8.1) (17). Region of the initial states of system $\tilde{S}_0$. (18). Segment AB (see Fig. 8.1). (19). Region of final states of system. (20). Triangle AOB (except for hypotenuse) (see Fig. 8.1). (21). Optimum initial state of system. (22). Optimum quantity of means x^*_1, isolated in first branch, and determined by it quantity of means $y^*_1 = Z_0 - x^*_1$, isolated in second branch. (23). Optimum control. (24). Optimum quantity of means over years, separating into branch I.

In the following presentation we shall everywhere follow overall diagram §7, no longer accompanying it by so/such comprehensive by explanations.

§9. Examples of the tasks about the distribution of resources/lifetimes.

For mastering the general solution of the task about the distribution of resources/lifetimes, given in the previous paragraph, it is useful to use it on the concrete/specific/actual material. Here we will consider two specific examples of general problem about the distribution of the resources/lifetimes, in each of which let us assign the completely specific form of the function f(x), g(y), $\varphi(x)$, $\psi(y)$, and let us bring each of the examples to the numerical result.

Example 1. Is planned/glided the work of two branches of production I and II for period m of years.

A quantity of means x, enclosed in branch I, gives in one year the income

$$f(x) = x^2 \qquad (9.1)$$

and due to this it is reduced to

$$\varphi(x) = 0.75x. \qquad (9.2)$$

A quantity of means y, enclosed in branch II, gives in one year the income

$$g(y) = 2y^2 \qquad (9.3)$$

and it is reduced to

$$\psi(y)=0.3y^{1)}. \qquad (9.4)$$

FOOTNOTE [1]. Unity the measurement of income and imbedded means must not be the same. The use/application of formulas of type (9.1) and (9.3) does not contradict the principles of dimensions, if means and income are expressed in completely specific units of measurement. ENDFOOTNOTE.

Page 67.

It is necessary to produce the distribution of service lives Z_0 between branches of the I and II to each year period being planned.

Solution. Conditional optimum control x_m^* on the latter/last step/pitch (quantity of means, isolated in branch I) is located as value x_m, with which it reaches maximum income at the latter/last step/pitch:

$$W_m^*(Z_{m-1})=\max_{0<x_m<Z_{m-1}}\{w_m(Z_{m-1},\ x_m)\}.$$

where

$$w_m(Z_{m-1},\ x_m)=x_m^2+2(Z_{m-1}-x_m)^2. \qquad (9.5)$$

The graph of function

$$w_m=w_m(Z_{m-1},\ x_m)$$

depending on argument x_m is represented with the given one Z_{m-1} by certain parabola (Fig. 9.1). The second derivative of function w_m of x_m is positive, and therefore parabola is converted concave-up. Maximum value can be reached only on the borders of gap/interval $(0, Z_{m-1})$[2].

FOOTNOTE [2]. Therefore has no sense to attempt to seek the maximum of function w_m, equating the derivative to zero. ENDFOOTNOTE.

In order to determine, on what precisely border, let us substitute into formula (9.5) $x_m=0$ and $x_m=Z_{m-1}$. We will obtain in the first case (when $x_m=0$)

$$w_m=2Z^2_{m-1},$$

in the second case (when $x_m=Z_{m-1}$)

$$w_m=Z^2_{m-1}.$$

The first value more than the second: consequently, independent of value Z_{m-1}, the maximum of income at the latter/last step/pitch reaches when $x_m=0$, i.e., conditional optimum control $x^*_m(Z_{m-1})$ does not depend on Z_{m-1} and it is always equal to zero, but this means that in the beginning of last year all available means it is necessary to pack into branch II.

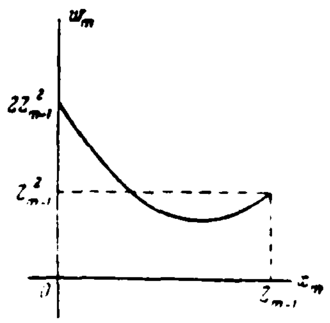


Fig. 9.1.

Page 68.

This is only logical, since income from this branch is more, but the expenditure of resources us no longer interests (following step/pitch it will not be).

During this optimum control last year will bring to us the income

$$W^*_m(Z_{m-1}) = 2Z^2_{m-1}.$$

Let us switch over to the distribution of resources to (m-1)-th year. Let we approach it with the supply of resources Z_{m-2}. Let us find the conditional maximum income in two last year:

$$W^*_{m-1,m}(Z_{m-2}) =$$
$$= \max_{0 \leq x_{m-1} \leq Z_{m-2}} \{x^2_{m-1} + 2(Z_{m-2} - x_{m-1})^2 + W^*_m(Z_{m-1})\}.$$

But

$$Z_{m-1} = 0.75x_{m-1} + 0.3(Z_{m-2} - x_{m-1}).$$

and consequently,

$$W^*_m(Z_{m-1}) = 2[0.75x_{m-1} + 0.3(Z_{m-2} - x_{m-1})]^2.$$

Hence we will obtain

$$W^*_{m-1,m}(Z_{m-2}) = \max_{0 \le x_{m-1} \le Z_{m-2}} \{x^2_{m-1} + 2(Z_{m-2} - x_{m-1})^2 + 2[0.75x_{m-1} + 0.3(Z_{m-2} - x_{m-1})]^2\}.$$

The expression in the curly brackets, briefly designated $W_{m-1,m}$, is again the polynomial of the second degree relatively x_{m-1} with the positive second derivative, and its graph - parabola with convexity downward, so that it is again necessary to trace to the maximum only the extreme points of interval (Fig. 9.2):

$$x_{m-1} = 0 \text{ (1) } x_{m-1} = Z_{m-2}.$$

Key: (1). and.

In the first case (when $x_{m-1} = 0$) we will obtain

$$W^*_{m-1,m} = 2Z^2_{m-2} + 2(0.3Z_{m-2})^2 = 2.180Z^2_{m-2};$$

in the second case (when $x_{m-1} = Z_{m-2}$)

$$W^*_{m-1,m} = Z^2_{m-2} + 2(0.75Z_{m-2})^2 = 2.125Z^2_{m-2}.$$

whence it is clear that the maximum again reaches when $x_{m-1}=0$ and is equal to $W^{*}_{m-1,m}(Z_{m-2})=2.180Z^{2}_{m-2}$. i.e., at the next-to-last step/pitch it is necessary all resources to pack into branch II.

Let us pass toward (m-2)-th to step/pitch. It is here necessary to maximize the polynomial of the second degree

$$W^{+}_{m-2,m-1,m}=x^{2}_{m-2}+ \\ +2(Z_{m-3}-x_{m-2})^{2}+ \\ +2.18[0.75x_{m-2}+ \\ +0.3(Z_{m-3}-x_{m-2})]^{2}.$$

The corresponding parabola (as on any of the steps/pitches) will be again converted concave-up. But this time maximum will be reached not on the left, but on the right border of section (Fig. 9.3). Actually/really, assuming/setting $x_{m-2}=0$, we will obtain

$$W^{+}_{m-2,m-1,m}=2Z^{2}_{m-3}+2.18(0.3Z_{m-3})^{2}\approx 2.20Z^{2}_{m-3},$$

but when $x_{m-2}=Z_{m-3}$

$$W^{+}_{m-2,m-1,m}=Z^{2}_{m-3}+2.18(0.75Z_{m-3})^{2}\approx 2.23Z^{2}_{m-3}.$$

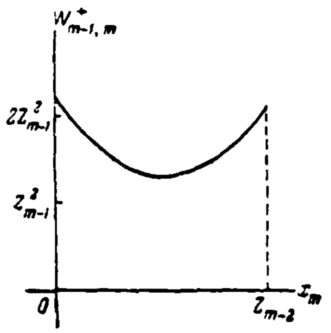


Fig. 9.2.

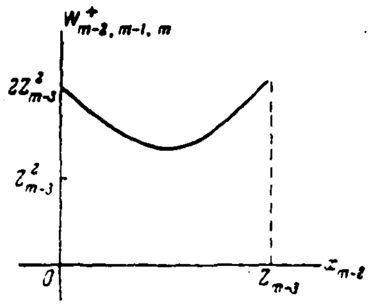


Fig. 9.3.

Page 70.

Consequently, conditional optimum control at (m-2)-th step/pitch will be

$$x^{*}_{m-2}(Z_{m-3}) = Z_{m-3},$$

i.e., on this step/pitch optimum control lies in the fact that all available resources to pack into branch I. In this case we will obtain the conditional maximum income

$$W^{*}_{m-2, m-1, m}(Z_{m-3}) \approx 2.23 Z^2_{m-3}.$$

It is obvious, in all following stages the maximum will be always reached as in Fig. 9.3, at the right end/lead of the segment. Actually/really, for i<m-2 function $W^{*}_{i, i+1, \ldots, m}$ will take the form

$$W^{+}_{i, i+1, \ldots, m} = \\ = x_i^2 + 2(Z_{i-1} - x_i)^2 + C[0.75x_i + 0.3(Z_{i-1} - x_i)]^2.$$

where coefficient C will be more than 2.18, since it with each step/pitch only increases. Therefore optimum conditional control to the very first step/pitch (inclusively) it will remain

$$x_i^*(Z_{i-1}) = Z_{i-1} \qquad (i = m-2,\ m-3,\ \ldots),$$

and conditional maximum income for all steps/pitches, beginning from the i-th, it will be

$$W^*_{i,\,i+1,\,\ldots,\,m}(Z_{i-1}) = Z^2_{i-1} + W^*_{i+1,\,\ldots,\,m}(0.75 Z_{i-1}).$$

Thus, optimum control is found: it lies in the fact that at all steps/pitches, except next-to-last and latter, to pack all resources into branch I, and at the latter/last steps/pitches to pack all resources into branch II. Let us note that this solution is obtained independently neither of a number of steps/pitches m nor of the initial supply of resources Z_0.

In order to visualize the type of optimum trajectory in the phase space, let us assign the concrete/specific/actual value of a number of steps/pitches m=5 (production process is planned/glided to 5 years).

Page 71.

Optimum trajectory is represented in Fig. 9.4. Optimum control process of resources consists of the following. To the first year all resources are packed into branch I and are reduced to 0.75 Z_0. By the second year - into the same branch I they are reduced to 0.56 Z_0 (there is no redistribution of resources, and therefore the second component/link of the second step/pitch vanishes). On the third year again all resources are packed into the same branch I and are reduced to 0.42 Z_0. On the fourth year the policy varies: occurs the redistribution of resources (inclined trajectory phase), they all are packed into branch II and are reduced to 0.13 Z_0. On the latter, the fifth, to year again all resources are packed into branch II; their remainder/residue at the end of the fifth year (and entire period) will be equal to 0.04 Z_0. During this distribution of resources in the five-year plan will be obtained the maximum income, equal to

$$W^* = 2.27 Z_0.$$

Of this example optimum control consisted of at each step/pitch packing of all resources either into one or into another branch. Always whether this will be thus? Now we will ascertain that not always. For this change the form of the function f(x) and g(y).

Example 2. Is planned/guided the activity of two branches of production with the I and II period to 5 years (m=5). The "functions of the expenditure of resources" $\varphi(x) = 0.75x$ and $\psi(y) = 0.3y$ the same as in the previous example, but the "function of the income" of f(X) and g(y) of the replacement by others:

$$f(x) = 1 - e^{-x}; \quad g(y) = 1 - e^{-2y}.$$

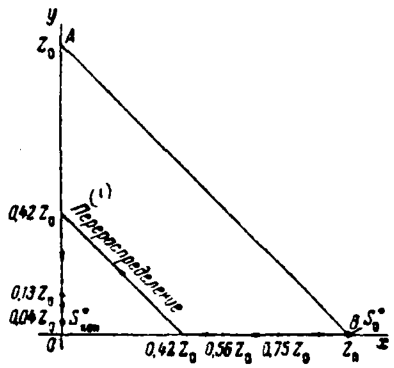


Fig. 9.4.

Key: (1). Redistribution.

Page 72.

It is necessary to distribute the available resources/lifetimes in size/dimension of $Z_0=2$ between branches I and II over the years.

Solution. In the previous example, in connection with the very simple form of the function $f(x)$ and $g(y)$, the solution was given in the analytical form; in this example to construct the analytical solution is difficult, and we will solve problem numerically. The meeting in the task functional dependences we will represent with the help of the graphs. Let at the beginning of the fifth year a quantity of resources be equal Z_4. In order to find conditional optimum control on the fifth step/piece $x^*_5(Z_4)$, it is necessary for each Z_4 to find the maximum of the function

$$W'_5 = w_5 = 1 - e^{-x_5} + 1 - e^{-2(Z_4 - x_5)} = \\ = 2 - [e^{-x_5} + e^{-2(Z_4 - x_5)}]. \quad (9.6)$$

With that fixed/recorded Z_4 this is - the function of argument x_5, convex upwards (Fig. 9.5). The maximum of this function (depending on value of Z_4) can be reached either within segment (0, Z_4) (as shown in Fig. 9.5a), or at his left end/lead (Fig. 9.5 b).

In order to find this maximum, let us differentiate expression (9.6) on x_5. If derivative becomes zero at certain point within the segment (0, Z_4), then at this point reaches maximum W; if outside - maximum reaches at x_5=0.

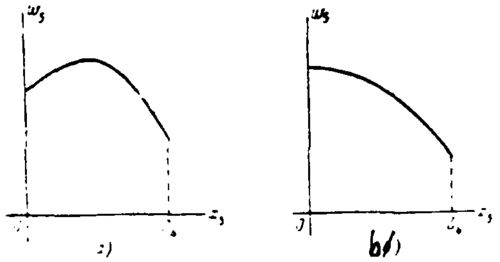


Fig. 9.5.

Page 73.

Differentiating (9.6) we have

$$\frac{\partial W_5}{\partial x_5} = e^{-x_5} - 2e^{-2(Z_4 - x_5)} = 0. \qquad (9.7)$$

At this step/pitch equation (9.7) to us still it is possible to solve in the literal form; at further steps/pitches analogous problems we will solve numerically. From (9.7) we have

$$-x_5 = \ln 2 - 2Z_4 + 2x_5; \quad x_5 = \frac{2Z_4 - \ln 2}{3}. \quad (9.8)$$

From expression (9.8) it follows that at $Z_4 > \ln 2/2 \approx 0.347$ the maximum reaches within the segment $(0, Z_4)$, at the point

$$x_5^*(Z_4) = \frac{2Z_4 - \ln 2}{3}. \qquad (9.9)$$

When $Z_4 < \frac{\ln 2}{2} \approx 0.347$ maximum reaches at the left end of the segment:

$$x_5^*(Z_4) = 0.$$

Thus, conditional optimum control on the fifth step/pitch is found:

$$x_5^*(Z_4)=\begin{cases} 0 & \overset{(1)}{\text{при}} \quad Z_4 \leqslant \frac{\ln 2}{2}, \\ \frac{2Z_4-\ln 2}{3} & \overset{(1)}{\text{при}} \quad Z_4 > \frac{\ln 2}{2}. \end{cases} \quad (9.10)$$

Key: (1). with.

Let us find conditional maximum income in the fifth year. It is equal to

$$W_5^*(Z_4)=2-\left\{e^{-x_5^*(Z_4)}+e^{-2[Z_4-x_5^*(Z_4)]}\right\}, \quad (9.11)$$

or, substituting (9.10) in (9.11),

$$W_5^*(Z_4)=\begin{cases} 1-e^{-2Z_4} & \overset{(1)}{\text{при}} \quad Z_4 \leqslant \frac{\ln 2}{2}, \\ 2-\frac{3}{2}\sqrt[3]{2}\,e^{-\frac{2}{3}Z_4} & \overset{(1)}{\text{при}} \quad Z_4 > \frac{\ln 2}{2}. \end{cases} \quad (9.12)$$

Key: (1). with.

Since for us it is necessary many times to compute value W^*_5, it will be convenient to construct its graph depending on Z_4 (Fig. 9.6).

Page 74.

On the same graph (but on other scale) let us depict the dependence of conditional optimum control at the fifth step/pitch x^*_5 on Z_4. With the construction of these two graphs are finished our all matters, connected with the fifth step/pitch. Subsequently, optimizing control at the fourth step/pitch, we will only enter into these graphs with different values of Z_4.

We pass to the fourth step/pitch. The task of its conditional optimization we will solve numerically, being assigned the series/row of values Z_3 (supply of the resources, which remained after the third step/pitch). In order not to make excess work, let us explain, within what limits can be found Z_3. Let us find the largest of possible of values Z_3. It will be achieved/reached, if at the first three steps/pitches all resources will be imbedded in branch I; in this case the supply of resources after three years will be equal to

$$Z_{3\,max} = Z_0 \cdot 0.75^3 = 0.844.$$

The smallest supply of resources corresponds to the case when all resources at three first steps/pitches are imbedded in branch II:

$$Z_{3\,min} = Z_0 \cdot 0.3^3 = 0.054.$$

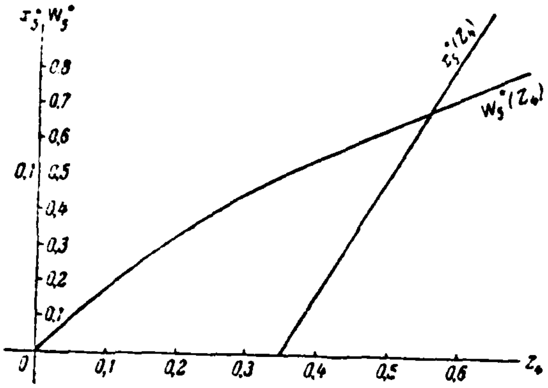


Fig. 9.6.

Page 75.

Thus, all possible values Z_3 are included in the section from 0.054 to 0.844. Let us assign in this section reference values of Z_3:

$$Z_3 = 0.1;\ 0.2;\ 0.3;\ 0.4;\ 0.5;\ 0.6;\ 0.7;\ 0.8 \quad (9.13)$$

and for each of them let us find conditional optimum control on the 4th step/pitch $x^*_4(Z_3)$ and conditional maximum income at two latter/last steps/pitches $W^*_{4,5}(Z_3)$. For this let us construct the series of the curves, which represent prize $W^+_{4,5}$ at two latter/last steps/pitches (during any control on the fourth and with the optimum - on the fifth):

$$W^+_{4,5} = w_4(Z_3,\ x_4) + W^*_5(0.75x_4 + 0.3(Z_3 - x_4)),$$

---

where

$$w_4(Z_3, x_4) = \\ = 2 - [e^{-x_4} + e^{-2(Z_3 - x_4)}].$$

and W^*_5 we find through the graph/curve Fig. 9.6, input into it with argument $Z_4 = 0.75x_4 + 0.3(Z_3 - x_4)$. The curves of dependence $W^*_{4,5}$ on x_4 (with the given one Z_3) are represented in Fig. 9.7. For each of these curves let us find point with the maximum ordinate and will mark by its small circle. The ordinate of this point for that corresponding to the curve Z_3 is conditional maximum prize at two latter/last steps/pitches $W^*_{4,5}(Z_3)$, and abscissa - conditional optimum control $x^*_4(Z_3)$. After determining these values for each value from (9.13), let us construct the graph/diagrams of dependences $W^*_{4,5}(Z_3)$ and $x^*_4(Z_3)$ (Fig. 9.8).

By the construction of these two curves we finished our calculations with two latter/last steps/pitches: all information about them is already included in of two curves of Fig. 9.8.

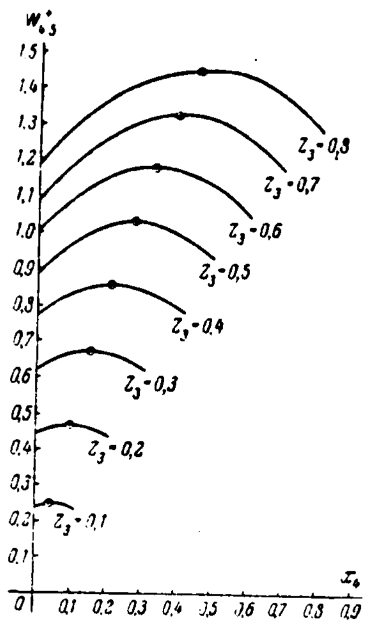


Fig. 3.7.

Page 76.

We pass to the third step/pitch. The region of the possible values Z_2 lies/rests between $2 \cdot 0.3^2 = 0.18$ and $2 \cdot 0.75^2 = 1.12$. We are assigned in this interval by the series/row of reference values Z_2:

$$Z_2 = 0.3;\ 0.5;\ 0.7;\ 0.9;\ 1.1$$

and for each of these values let us compute income on the third step/pitch depending on control x_3 at this step/pitch according to the formula

$$w_3(Z_2,\ x_3) = 2 - [e^{-x_3} + e^{-2(Z_2 - x_3)}].$$

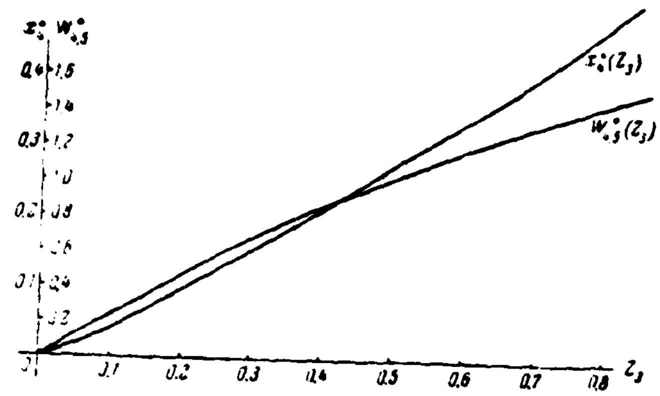
Then let us adjoin to it the already optimized income at fourth and fifth steps/pitches $W^*_{4,5}(Z_4)$, which we will determine on the graph/curve Fig. 9.8, entering it with the value

$$Z_4 = 0.75x_3 + 0.3(Z_2 - x_3),$$

and we will obtain the value

$$W^*_{3,4,5} = w_3(Z_2, x_3) + W^*_{4,5}(0.75x_3 + 0.3(Z_2 - x_3)),$$

for which let us again construct the graph/diagrams of dependence on x_3 with that fixed/recorded Z_2 (Fig. 9.9). For each of these curves let us again find the maximum (in the figure it is noted by small circle) and after this will construct the dependence of the conditional optimum control at the third step/pitch x^*_3 and of the corresponding to condition maximum income at three latter/last steps/pitches $W^*_{3,4,5}$ on Z_2 (Fig. 9.10).



x_3°, W_{*3}°
0.4
0.3
0.2
0.1
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Z_3
$x_3^\circ(Z_3)$
$W_{*3}^\circ(Z_3)$

Fig. 3.8.

Page 77.

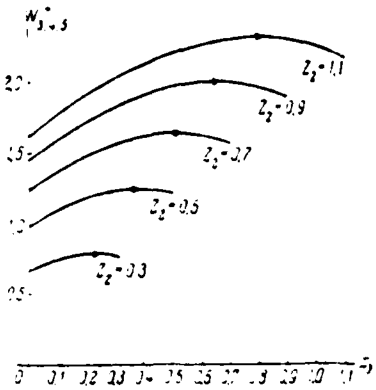


Fig. 9.9.

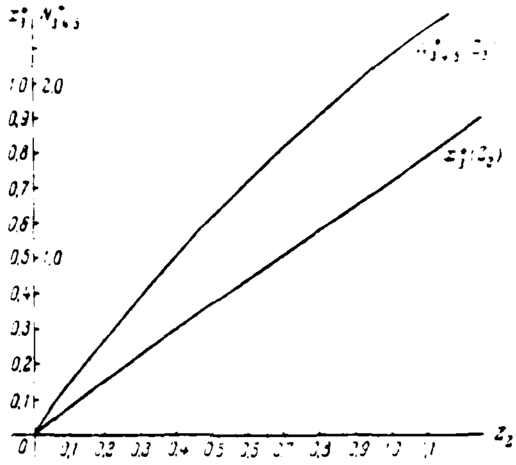


Fig. 9.10.

Page 78.

Analogously is solved the problem of the conditional optimization of the second step/stage: are varied values Z_1 from $2 \cdot 0.3 = 0.6$ to $2 \cdot 0.75 = 1.5$:

$$Z_1 = 0.6;\ 0.9;\ 1.2;\ 1.5.$$

Income at the second step/stage will be

$$w_2(Z_1, x_2) = 2 - [e^{-x_2} + e^{-2(Z_1 - x_2)}].$$

To it is adjoined the conditional maximum income $W^*_{3.4.5}$ on the graph/curve Fig. 9.10 with the input

$$Z_2 = 0.75x_2 + 0.3(Z_1 - x_2);$$

it is obtained value $W^+_{2,3,4,5}$, for which again they are constructed graphs (Fig. 9.11). On each curve is located the maximum and are constructed two curves: $x^*_2(Z_1)$ and $W^*_{2.3.4.5}(Z_1)$ (Fig. 9.12).

It remained to plan the first step/pitch. This - already more easy problem, since value Z_0, with which we begin this step/pitch, it is accurately known (Z_0=2) and it must not be varied. Therefore for the first step/pitch is constructed only one curve dependence $W^+_{1,2,3,4,5}$ on x_1 (Fig. 9.13), where

$$W^+_{1,2,3,4,5} = w_1(Z_0, x_1) + W^*_{2,3,4,5}(Z_1) = \\ = 2 - \left[e^{-x_1} + e^{-2\cdot(Z_0 - x_1)}\right] + W^*_{2,3,4,5}(Z_1),$$

and latter/last term is located through the graph/curve Fig. 9.12 with

$$Z_1 = 0.75x_1 + 0.3(Z_0 - x_1),$$

where Z_0=2.

Determining in the unique curve of Fig. 9.13 maximum, we find (no longer conditional) optimum control on the first step/pitch x^*_1=1.6 and corresponding maximum income in all five years

$$W^* = W^*_{1,2,3,4,5} = 4.35.$$

Page 79.

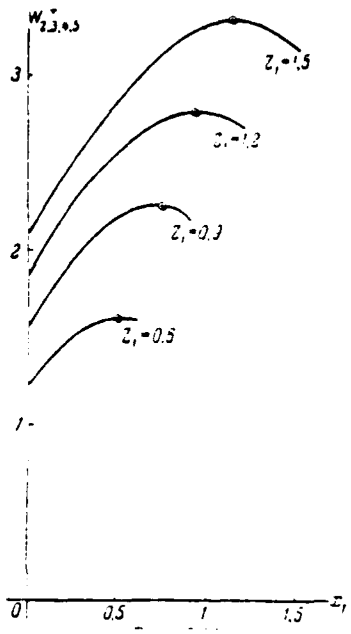


Fig. 9.11.

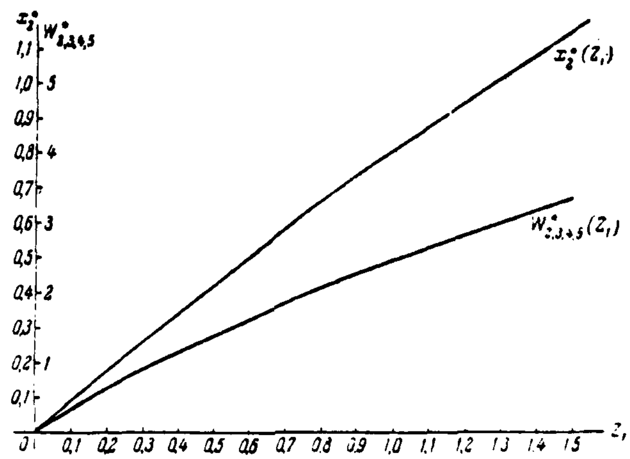


Fig. 9.12.

Page 80.

After this, as always in the method of dynamic programming, it is necessary to construct complete optimum control

$$X^* = (x_1^*,\ x_2^*,\ x_3^*,\ x_4^*,\ x_5^*),$$

going in the opposite direction: from the first step/pitch toward the fifth.

Knowing optimum control on the first step/pitch

$$x_1^* = 1.60.$$

we find the corresponding to it supply of resources toward the end of the first step/pitch:

$$Z_1^* = 0.75x_1^* + 0.3\,(Z_0 - x_1^*) = 1.32.$$

Entering with this value of Z_1 into graph $x^*_2(Z)$ (see Fig. 9.12), we find optimum control on the second step/pitch:

$$x_2^* = 1.02.$$

The remainder/residue of resources toward the end of the second step/pitch will be

$$Z_2^* = 0.75x_2^* + 0.3\,(Z_1^* - x_2^*) = 0.86.$$

With this value of Z_2 we enter into graph $x^*_3(Z_2)$ (cm Fig. 9.10) and find optimum control on the third step/pitch:

$$x_3^* = 0.62.$$

The remainder/residue of resources after the third step/pitch will be $Z^*_3=0.75x^*_3+0.3(Z^*_2-x^*_3)=0.54$. Through the graph/curve Fig. 9.8 we find optimum control on the fourth step/pitch

$$x_4^* = 0.30.$$

After the fourth step/pitch the remainder/residue is equal to

$$Z_4^* = 0.75x_4^* + 0.3\,(Z_3^* - x_4) = 0.30.$$

With this value of Z_4 we enter into graph $x^*_5(Z_4)$ (see Fig. 9.6) and find optimum control on the latter/last step/pitch

$$x_5^* = 0.$$

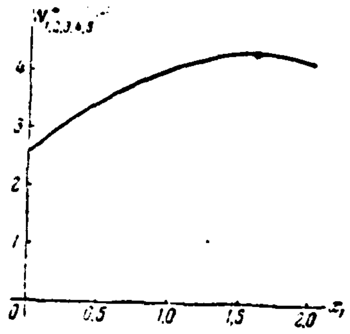


Fig. 9.13.

END SECTION.

Page 31.

Thus, planning/gliding process is completed. Is found the optimum control, which indicates, how many resources from the available supply $Z_0=2$ it is necessary to pack into branch I over the years:

$$X^*=(1.60;\ 1.02;\ 0.62;\ 0.30;\ 0).$$

Taking into account that the supplies of the resources before beginning each year are known:

$$Z_0=2;\quad Z_1^*=1.32;\quad Z_2^*=0.86;\quad Z_3^*=0.54;\quad Z_4^*=0.30,$$

we automatically obtain quantities of resources, packed over the years into branch II:

$$y_1^*=Z_0-x_1^*=0.40;\quad y_2^*=Z_1^*-x_2^*=0.30;$$
$$y_3^*=Z_2^*-x_3^*=0.24;\quad y_4^*=Z_3^*-x_4^*=0.24;$$
$$y_5^*=Z_4^*-x_5^*=0.30.$$

Thus, it is possible to formulate the following recommendations regarding the optimum distribution of resources. From the available in the beginning period of the supply of resources $Z_0=2$ and remaining resources at the end of each year it is necessary to pack over the years in branch the I and II following sums:

Год	1-й	2-й	3-й	4-й	5-й
I	1,60	1,02	0,62	0,30	0
II	0,40	0,30	0,24	0,24	0,30

During this planning/gliding will be obtained maximal return in 5 years, the equal to

$$W^{*}_{1,2,3,4,5} = 4.35.$$

Remainder/residue of resources at the end of the period will be equal to

$$0.3 \cdot 0{,}30 = 0.09.$$

Fig. 9.14 depicts the optimum trajectory in the phase space, which corresponds to this distribution of resources. Point S^*_0 on the hypotenuse of triangle AOB represents the optimum initial distribution of resources with the sharp predominance to the side of branch I. The first component/link of broken line corresponds to the expenditure of resources in the first year.

Page 82.

The following components/links are joined they pair-wise and represent redistribution and expenditure of resources on the 2nd, 3rd, 4th and 5th years. Latter/last component/link lies/rests on axis Oy; this means that on the 5th year of production process all resources are packed into branch II. Point $S^{*}_{кон}$ represents the remainder/residue of resources Δ^*_5=0.09, which is obtained during the optimum planning/gliding.

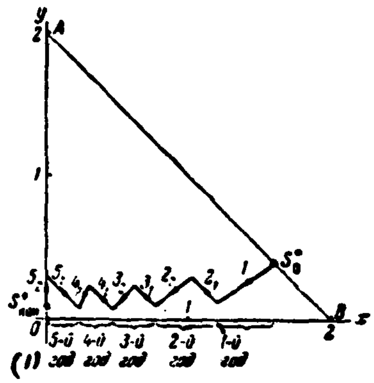


Fig. 9.14.

Key: (1). year.

§ 10. Modifications of the task about the distribution of resources/lifetimes.

the examined in two previous paragraphs task about the distribution of resources/lifetimes has many modifications. Some of them comparatively differ little from the simplest task, examined into § 8; others so differ from it in their verbal/literary setting, which is sometimes difficult to discover in them general/common/total features. In this paragraph and in those following (§§ 11.12) we will consider the series/row of the versions of similar tasks.

a. Distribution of resources/lifetimes in heterogeneous stages. In the task § 8 stages (steps/pitches) were "uniform" in the sense that resources x and y, intended respectively in branch I and II, in any stage gave one and the same income and were reduced in an identical way independent of the number of stage.

Page 83.

The natural generalization of this simplest task is the case when income and loss/depreciation of resources in different stages are dissimilar: resources x, y, intended in branch I and II, give on the i-th income $f_i(x)$, $g_i(y)$ and they are reduced to $\varphi_i(x) \leqslant x$, $\psi_i(y) \leqslant y$.

How can arise this heterogeneity? By different methods. For example, profitableness can depend on the common level of the development of production, achieved/reached to the defined period; or the condition of production (as, let us say, in the agriculture) they can depend on season.

For the solution of the problem of distributing the resources/lifetimes by the method of dynamic programming this circumstance - uniformity or heterogeneity stage- is completely unessential. Since the problem of the optimization of control nevertheless is solved in stages, it is completely unimportant, are identical functions $f_i(x)$, $g_i(y)$, $\varphi_i(x)$, $\psi_i(y)$ in the different stages or they are different.

The overall diagram of the solution is reduced to the consecutive use/application of the following formulas for the conditional optimum income in several latter/last stages:

$$W_m^*(Z_{m-1}) = \max_{0<x_m<Z_{m-1}} \{f_m(x_m) + g_m(Z_{m-1} - x_m)\};$$

$$W_{m-1,m}^*(Z_{m-2}) = \max_{0<x_{m-1}<Z_{m-2}} \{f_{m-1}(x_{m-1}) + g_{m-1}(Z_{m-2} - x_{m-1}) + W_m^*(\varphi_{m-1}(x_{m-1}) + \psi_{m-1}(Z_{m-2} - x_{m-1}))\};$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$W_{i,i+1,\ldots,m}^*(Z_{i-1}) = \max_{0<x_i<Z_{i-1}} \{f_i(x_i) + g_i(Z_{i-1} - x_i) + W_{i+1,\ldots,m}^*(\varphi_i(x_i) + \psi_i(Z_{i-1} - x_i))\}.$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

with the incidental definition of the conditional optimum controls:

$$x_m^*(Z_{m-1}),\quad x_{m-1}^*(Z_{m-2}),\ \ldots,\ x_1^*(Z_0).$$

After this, as always, is constructed optimum control, beginning from the first stage and ending with the latter. In this construction of there is no difference with the case of uniform stages.

Page 34.

5. Task about redundancy of resources/lifetimes. Task is placed as follows. There is only one branch of production and certain supply of resources Z_0, which can be packed into the production not wholly, but partially be reserved. Being in the production in the i stage, a quantity of resources x imbedded gives income $f_i(x)$ and and it is reduced to $\varphi_i(x) < x$. It is necessary to rationally distribute the available and remaining resources in m stages in order to become maximum aggregate profit W.

It is not difficult to ascertain that this task is reduced to previous. Actually/really, the reserved resources can be considered "imbedded" in certain fictitious "second branch" of production, in which the resources are not expended, but also they do not give the income:

$$g_i(y)=0;\quad \psi_i(y)=y \qquad (i=1,\ 2,\ \ldots,\ m).$$

Taking into account this condition problem is solved in exactly the same way just as the task of distributing the resources/lifetimes.

The trajectory of point S, which represents the state of system in the phase space, will take the form, represented in Fig. 10.1. The sections of the "redistribution of resources" will be, as before they are parallel to line AB, while the sections of the "consumption of resources" - are parallel to the axis of abscissas and are directed to the left. The latter/last component/link of broken line will always lie/rest on the axis of abscissas, since further redundancy of resources a sense does not have.

Let us consider a special case of the task about the redundancy when in all stages

$$\varphi_i(x)=0,$$

i.e. the imbedded resources are expended/consumed by pillar. Then the task of the redundancy of resources is reduced to finding of the maximum of the following function m of arguments:

$$W = \sum_{i=1}^{m} f_i(x_i). \qquad (10.1)$$

where $x_1, x_2, \ldots, x_m$ limited by the conditions

$$\sum_{i=1}^{m} x_i \leqslant Z_0; \qquad (10.2)$$

$$x_i \geqslant 0. \qquad (10.3)$$

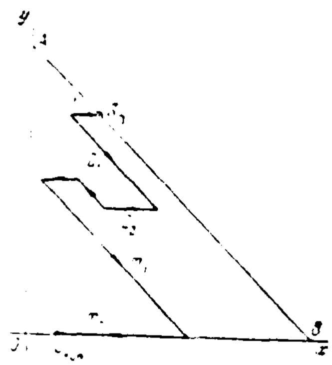


Fig. 10.1.

Page 85.

If we income $f_i(x)$ (as this logical assume) is the nondecreasing function of the imbedded resources x, then the sign of equality in formula (10.2) can be rejected/thrown, since under these conditions to expend/consume not all resources, but only their part is disadvantageous.

The trajectory of point S in the phase space will appear, as shown in Fig. 10.2 - each horizontal section reaches the axis of ordinates.

Let us do some observations about the method of the solution of problem. Above we saw that she was reduced to the determination of the maximum of function (10.1). It can seem that thereby the task is simplified, according to this impression illusory. Indeed generally the task of finding the maximum of the function of many arguments is not an easy one. Let us recall (see § 1) that any task of optimum control is always reduced to finding the maximum (minimum) of the function of many arguments, and precisely in order to avoid the connected arguments, and precisely in order to avoid connected with this difficulties, we resort to the method of dynamic programming. After giving here formula (10.1), we did not intend to facilitate the task of dynamic programming, after reducing it to the task of the determination of the maximum of function (10.1). On the contrary, for the solution of the problem of the determination of the maximum (minimum) of function of type (10.1) with conditions (10.2) and (10.3) (wherever this task not it arose), can prove to be most adequate/approaching precisely the method of dynamic programming. By the use/application of this method in this case we bring the multidimensional task of finding the maximum of the function of many variable/alternating to the repeated determination of the maximum of the function of one variable/alternating, which is considerably easier.

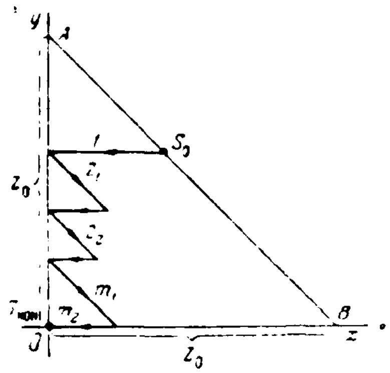


Fig. 10.2.

Page 86.

Let us note, however, that some simplest cases of the task of the redundancy of resources admit elementary solution, also, without the use/application of a method of dynamic programming. To them belongs, for example, simplest case when the "function of income" in all stages is one and the same:

$$f_1(x) = f_2(x) = \ldots = f_m(x) = f(x).$$

moreover resources in each stage are expended/consumed completely:

$$\varphi_1(x) = \varphi_2(x) = \ldots = \varphi_m(x) = 0.$$

it is possible to demonstrate that if function f(x) - function monotonically increasing and is convex upward (Fig. 10.3), then the maximum of expression (10.1) reaches, when resources are divided into equal parts between all stages:

$$x_1^* = x_2^* = \dots$$
$$\dots = x_m^* = \frac{Z_0}{m}.$$

3. Task about the distribution of the resources/lifetimes between several (more than of two) branches. The task about the distribution of resources/lifetimes allows/assumes generalization to the case when resources are distributed not between two, but between k branches:

$$\mathrm{I},\ \mathrm{II},\ \dots,\ (k),$$

moreover for each (j-th) branch they are preset: the "function of income"

$$f_i^{(j)}(x),$$

expressing the income, given by a quantity of resources x, imbedded in the j-th branch at the i step/pitch, and the "function of expenditure"

$$\varphi_i^{(j)}(x) \leqslant (x),$$

showing, to which value decreases a quantity of resources x, imbedded in the j-th branch at the i step/pitch.

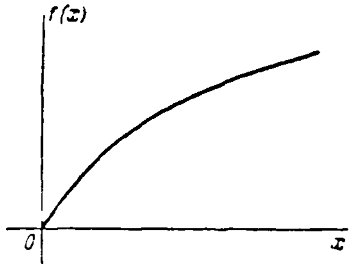


Fig. 10.3.

Page 87.

Let us construct for this task phase space. In the case of distributing the resources according to two branches such phase space was triangle AOB (see Fig. 8.1, 8.2, etc.). For the case of several branches it is possible as the phase space to consider the multidimensional generalization of triangle (which is conventionally designated as "simplex"), namely the point set of k- graduated space, which satisfy the conditions:

$$\left.\begin{array}{l}\sum_{j} x^{(j)} \leqslant Z_0; \quad x^{(j)} \geqslant 0 \\ ((j = \mathrm{I}, \mathrm{II}, \ldots, (k)).)\end{array}\right\} (10.4)$$

In the case of the space of three measurements (which corresponds to the distribution of resources according to three branches) simplex will take the form of tetrahedron ABCD (Fig. 10.4) whose three edges/fins, that converge in the beginning of coordinates, are equal to Z_0. The process of distributing the resources, as in the two-dimensional case, it can be divided into the components/links, which correspond to the "redistribution of resources" and to the "expenditure of resources", moreover on the first components/links point S moves on the plane, parallel ABC, and on the second it moves, receding from plane ABC into the depth of simplex.

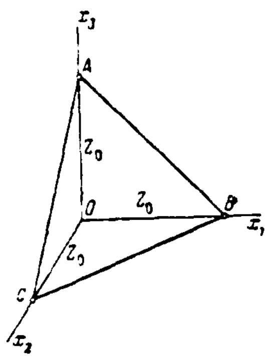


Fig. 10.4.

§ 11. Task about the distribution of resources/lifetimes with the enclosure of incomes into the production.

Until now, in all tasks examined about the distribution of resources/lifetimes we examined the "income", yielded by production, completely independent of the distributed basic means (it even could be expressed in other unity, for example resources/lifetimes - in the man-hours, and income - in the rubles or in the meters of fabric).

Page 88.

In this paragraph we will consider that case when income can (in full or in part) be packed into the production together with the basic means. For this it goes without saying the income and basic means must be given to one equivalent (for example, to the money).

Depending on situation this task can be placed differently, with the different criteria W. For example, it is possible to pack into the production entire income or its only certain of fraction/portion. It is possible to seek such control which ensures maximum total net income from m stages. It is possible to seek such control which converts into the maximum the total sum of resources (switching on income and preserved basic means) after m stages. Are possible other formulations of the problem. Here we will show the diagram of the solution by the method of the dynamic programming of several simplest tasks of such type.

a. Let us consider case when income is packed into production completely, moreover is maximized sum of all means (basic means plus income) after m stage.

In this case criterion W is the sum of all resources, which were preserved in both branches after the m stage, plus the income, given by both branches in this stage.

The criterion W in question is a special case of the additive criterion: it entire is acquired in the last stage, i.e., $W = w_m$, and in all previous stages its increases w_i are equal to zero.

Since all resources (and the remainder/residue of bases, and income) are packed into the production and are considered in criterion W on the equal bases/bases, then to us to unnecessarily here build-in separately the "functions of income" $f_i(x)$, $g_i(y)$ and the "functions of expenditure" $\varphi_i(x)$, $\psi_i(y)$ and is sufficient to introduce two functions

$$F_i(x), \quad G_i(y), \qquad (11.1)$$

showing, how many resources (remainder/residue of bases plus income) we will have at the end of the i stage, after putting in the beginning of this stage a quantity of resources x into the first branch and u the secondly. Let us name functions $F_i(x)$, $G_i(y)$ the "functions of a change in the resources" in the i stage. Let us note that is possible any of the relationships/ratios:

$$F_i(x) < x; \quad F_i(x) \geq x; \quad F_i(x) > x$$

(it is analogous for $G_i(y)$).

Page 39.

let us consider the phase space, which corresponds to this task (Fig. 11.1). Such space will be no longer triangle AOB (as in the tasks without the enclosure of incomes), but entire first quadrant xOy (resources can not only be reduced, but also increase). Trajectory as before consists of the series/row of components/links; to each stage (except the first) corresponds the pair of the components/links: the first - "redistribution of resources", when point S is moved in parallel AB; the second - "expenditure and the acquisition of resources", during which point S can move in any direction. In contrast to all previous examples, here obtaining the "final income" of W is connected only with one, latter itself, component/link m_2, which in Fig. 11.1 is isolated with heavy arrow.

In this case the value of criterion W is directly evident on the drawing - this is the sum of abscissa and ordinate of point $S_{кон}$ corresponding to final state system. Thus, the task of optimum control can be formulated so: to select this trajectory of point in the phase space in order to deduce it as a result of the m step/pitch for straight line $A_{кон}B_{кон}$ parallel AB and distant behind the origin of coordinates so far, as soon as this will be possibly. The value of criterion W is represented as the segment, intercepted/detached for each of the axes of straight line $A_{кон}B_{кон}$.

Let us construct the diagram of the solution of this problem by the method of dynamic programming without the comprehensive verbal/literary explanations, since the entity of method is sufficiently clear from previous. During function $F_i(x)$, $G_i(y)$ thus far we will superimpose no limitations.

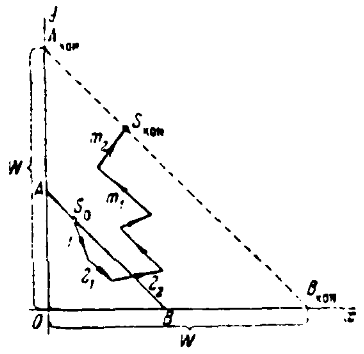


Fig. 11.1.

Page 90.

1. We fix/record issue (m-1)-th step/pitch (preserved resource plus income) Z_{m-1}. Conditional optimum control $x^*_m(Z_{m-1})$ - that with which will be maximum a total quantity of resources (basic means of plus return), after the m step/pitch

$$w_m(Z_{m-1}) = Z_m(Z_{m-1}). \qquad (11.2)$$

But, taking into account formulas (11.1), it is possible to write

$$w_m(Z_{m-1}) = F_m(x_m) + G_m(Z_{m-1} - x_m).$$

Conditional optimum control on m step/pitch $x^*_m(Z_{m-1})$ will be located from condition

$$W^*_m(Z_{m-1}) = \max_{0 < x_m < Z_{m-1}} \{F_m(x_m) + G_m(Z_{m-1} - x_m)\}. \qquad (11.3)$$

2. We fix/record issue (m-2) -th step/pitch Z_{m-2}. Conditional optimum control $x^*_{m-1}(Z_{m-2})$ is found from the condition

$$W^*_{m-1,m}(Z_{m-2}) = \max_{0 \leq x_{m-1} \leq Z_{m-2}} \{W^*_m(F_{m-1}(x_{m-1}) + G_{m-1}(Z_{m-2} - x_{m-1}))\} \quad (11.4)$$

and so forth.

3. We fix/record Z_{i-1}. Conditional optimum control $x^*_i(Z_{i-1})$ is found from the condition

$$W^*_{i,i+1,\ldots,m}(Z_{i-1}) = \max_{0 \leq x_i \leq Z_{i-1}} \{W^*_{i+1,\ldots,m}(F_i(x_i) + G_i(Z_{i-1} - x_i))\} \quad (11.5)$$

and so forth.

4. Optimum control at first step/pitch x^*_1 and maximum value of prize W are found from condition

$$W^* = W^*_{1,2,\ldots,m} = \max_{0 \leq x_1 \leq Z_0} \{W^*_{2,\ldots,m}(F_1(x_1) + G_1(Z_0 - x_1))\}.$$

5. Issue of the first step/pitch during the optimum control:

$$Z^*_1 = F_1(x^*_1) + G_1(Z_0 - x^*_1).$$

optimum control at the second step/pitch:

$$x^*_2 = x^*_2(Z^*_1).$$

Issue of the second step/pitch with the optimum of the controls:

$$Z_2^* = F_2(x_2^*) + G_2(Z_1^* - x_2^*)$$

and so forth to the latter/last step/pitch.

Page 91.

Is such the diagram of the solution of problem by the method of dynamic programming with any form of the function of a change in resources $F_i(x)$, $G_i(y)$. However, if we on these functions superimpose some (very natural) limitations, this diagram can be highly simplified.

Let us assume that all functions

$$F_i(x), \quad G_i(y) \qquad (i = 1, \ldots m)$$

are the nondecreasing functions of their arguments (i.e. that with an increase in the quantity of invested resources the sum of income and remaining resources toward the end of the stage it cannot decrease).

Let us show that under these conditions the maximum prize at the latter/last step/pitch is nondecreasing function from the issue of each step/pitch (sum of resources in its end/lead).

Let us consider maximum prize when the sum of resources (remainder/residue plus income) at the end (i-1) -th stage is equal to Z_{i-1}. Since prize is acquired only in the latter/last stage, the nevertheless, to examine this prize for entire process, either only for the latter/last stage, or for all stages, beginning from the i-th. Let us select the latter: we will examine maximum prize for all stages, beginning from the i-th as function from Z_{i-1}, designating it, as always

$$W^{*}_{i,\, i+1,\, \ldots,\, m}(Z_{i-1}).$$

Let us demonstrate that this function not decreasing. Proof we will conduct by full/total/complete induction, but not from i to $i+1$, as this is done usually, but on the contrary, from i+1 to i (in accordance with the "reverse" course of the process of dynamic programming).

Let us assume that the proven property is correct for i+1, i.e., the function

$$W^{*}_{i+1,\, \ldots,\, m}(Z_{i})$$

is the nondecreasing function of its argument Z_i (this it means: the greater the resources, switching on income and basic means, it was preserved to the issue of the i step/pitch, the greater there will be the income at the end). Let us demonstrate that then by nondecreasing function it will be and

$$W^{*}_{i,\, i+1,\, \ldots,\, m}(Z_{i-1}).$$

Page 92.

Actually/really, according to formula (11.5), $W^*_{i,\,i+1,\,\dots,\,m}(Z_{i-1})$ is the maximum of the expression

$$W^*_{i+1,\,\dots,\,m}(F_i(x_i)+G_i(Z_{i-1}-x_i)). \qquad (11.6)$$

Let us show that expression (11.6) is nondecreasing function Z_{i-1}; then it will be it is clear that as its maximum value $W^*_{i,\,i+1,\,\dots,\,m}(Z_{i-1})$ with increase Z_{i-1} decrease cannot.

Let us fix some value Z_{i-1}. let for this value Z_{i-1} expression (11.6) reach maximum in x_i, equal to $W^*_{i,\,i+1,\,\dots,\,m}(Z_{i-1})$ during the specific control (distribution of resources) x_i^*. Let us give now to value Z_{i-1} certain positive increase ΔZ_i. was formed certain surplus of resources, which we can distribute between branches I and II, after increasing a quantity of resources, imbedded either in one or in another branch, or into that and another immediately. Since function $F_i(x)$, $G_i(y)$ not decreasing, the from this "addition" of resources each of the components/terms/addends under the sign of function (11.6) can only be increased; it means, and their sum can only be increased, but not made is less.

What in this case will be with function (11.6)? According to assumption these are - function is; it means, and with increase Z_{i-1} it be reduced cannot. Thus, transition from i+1 to i is proved.

Let us show now that our property is correct for i+1=m, i.e., for the latter/last step/pitch. This is proven very simply. Prize at the latter/last step/pitch during the optimum control is the maximum of the expression

$$F_m(x_m) + G_m(Z_{m-1} - x_m)$$

and, naturally, to eat nondecreasing function from Z_{m-1} (this recently it was shown for any value of i, and also, therefore, for i=m). Thus, $W^*_m(Z_{m-1})$ is nondecreasing function Z_{m-1}, and means, according to the principle of full/total/complete induction, and any of the prizes $W^*_{i,i+1,\dots,m}(Z_{i-1})$ - nondecreasing function, QED.

Page 93.

From that proved escape/ensue very simple recommendations regarding the optimum control. Actually/really, if final prize W^*_m is nondecreasing function from the total sum of resources, realized on the issue of each step/pitch, then optimum control lies in the fact that on the issue of each step/pitch individually to obtain the maximum value of this sum of resources.

This means that in this special case the "interests" of operation as a whole coincide with the "interests" of each single step/pitch. The rational planning/gliding of entire operation is reduced to optimize each step/pitch individually, without worrying about the others.

This special feature/peculiarity leads to the fact that the process of producing the optimum control strongly is simplified. Actually/really, greater there is no necessity to fix/record the results of each previous step/pitch and to draw entire chain/network of conditional optimum controls from the latter/last step/pitch toward the first. It is possible to directly optimize step by step from the beginning toward the end. At the first step/pitch to take such control $x_1=x^*_1$, during which is converted into the maximum the sum of resources Z_1:

$$Z_1^* = \max_{0 < x_1 < Z_0} \{F_1(x_1) + G_1(Z_0 - x_1)\};$$

on the second - the control $x_2=x^*_2$, during which it is converted into maximum Z_2:

$$Z_2^* = \max_{0 < x_2 < Z_1^*} \{F_2(x_2) + G_2(Z_1^* - x_1)\}$$

and so forth to the end/lead.

Thus, with nondecreasing functions $F_i(x)$, $G_i(y)$ stated by us problem of the exterior only takes the form of task of dynamic programming, and actually - it is much simpler it.

Similar "degenerate" tasks of the dynamic programming where the optimum control lies in the fact that to optimize each etnane, without worrying about the others, frequently they are encountered in practice. If, without having focused attention on this special feature/peculiarity, to solve them nevertheless by the method of dynamic programming, the solution it goes without saying will be obtained accurate, but will require many times more time, than it is necessary.

Page 94.

Let us do one additional observation. At first glance it can seem that the superimposed during function $F_i(x)$, $G_i(y)$ condition - so that they would be nondiminishing - is satisfied in all in practice conceivable cases. However, it is possible to give the practical tasks, in which it is not implemented. Let us consider, for example, the case, when one of the "branches" of production is storage of the perishable goods (vegetables) on the storage. This branch yields only the losses, connected with the losses of goods during their storage. Let us designate $F_i(x) < x$ the value of commodities, which were being stocked, at the end of the i-th stage, if in the beginning of stage it was x. Always whether this function will be monotone? No, not always. It is possible to visualize such situation when with the overload of storage of more than certain critical value function F(x) begins to decrease (for example, due to deterioration in storage conditions). In similar cases is necessary to solve problem the cverall diagram of dynamic programming as this was shown above.

b. Let us consider case when into production as before is packed entire income, but criterion is net income in m stage (preserved basic means are not considered).

Let be preset to the "function of income" $f_i(x)$, $g_i(y)$ and to the "function of expenditure" $\varphi_i(x)$, $\psi_i(y)$ for each stage (i=1, ..., m).

Let us show that if function $f_m(x)$, $g_m(y)$ - the "function of income" in the latter/last stage - not decreasing, then task is reduced to examined in point/item A, namely to the maximization of total prize (remaining resource plus income) afterward (m-1) -th stage. Actually/really, conditional maximum prize at the latter/last step/pitch will be

$$W^*_m(Z_{m-1}) = \max_{0 < x_m < Z_{m-1}} \{f_m(x) + g_m(Z_{m-1} - x_m)\}. \quad (11.7)$$

It is possible to demonstrate (analogously how it was done in point/item a) that function $W^*_m(Z_{m-1})$ is the nondecreasing function of its argument, and its maximum reaches when Z_{m-1} it reaches its maximum value. Thus, for the determination of optimum control is sufficient to solve task "a" for first m-1 steps/pitches with the functions of a change in the resources

$$F_i(x) = f_i(x) + \varphi_i(x),$$
$$G_i(y) = g_i(y) + \psi_i(y)$$

and then to separately find optimum control on the m step/pitch, on the basis of formula (11.7).

Page 95.

If functions $F_i(x), G_i(y)$ $(i = 1, \ldots, m-1)$ will also be nondecreasing, then task, as in the preceding case, it will prove to be degenerate.

If functions $f_m(x), g_m(y)$ are not nondecreasing, then reducing to task "a" becomes already impossible and, is necessary to resort to the overall diagram of dynamic programming. To reader one should as the useful exercise sketch this diagram.

c. Let us consider case when income, obtained in each stage, is packed into production not completely, but partially, moreover is maximized full/total/complete net income in all stages plus remainder/residue of resources after m stage.

In this task, as in the ordinary task the distributions of resources/lifetimes, must be preset to the "function of income"

$$f_l(x),\quad g_l(y)\quad (l=1,\ 2,\ \ldots,\ m)$$

and the "function of expenditure"

$$\varphi_l(y)\leqslant x;\quad \psi_l(y)\leqslant y\quad (l=1,\ 2,\ \ldots,\ m).$$

Furthermore, must be preset to the "function of enclosure"

$$R_l(\xi)\leqslant \xi\quad (l=1,\ \ldots,\ m-1),$$

those showing, what part of income ξ obtained in the i stage, is packed into the production of the following, (i+1) -th, stage.

As the phase space let us consider no longer first quadrant xOy of plane, but first octant $xOy\xi$ of three-dimensional space (Fig. 11.2). Along the axes Ox and OY as before are plotted/deposited the resources, which are located in branches I and II; along the axis $O\xi$ - total income, yielded by both branches. Region $\tilde{S}_0$ of the initial states of system - as before hypotenuse AB of triangle AOB in plane xOy. All stages, except the first, are subdivided into two components/links: on the first component/link the resources (preserved in both branches plus the specific part of the income of the previous stage) are redistributed between the branches; on the second component/link occurs the expenditure of resources and the acquisition of income. Fig. 11.2 shows two stages: the first consists only of one component/link, the second - of two.

Page 96.

Let us consider values $x_1'=\varphi_1(x_1)\leqslant x_1$; $y_1'=\psi_1(y_1)\leqslant y_1$ - the resources, which were preserved in branches the I and II toward the end first stage where x_1, y_1 - coordinates of point S_0 - resources, imbedded in branch I and II during the first stage; $\xi_1=f_1(x_1)+g_1(y_1)$ - income, brought by both branches during the first stage. During the first stage point S, which represents the state of system, shifts from the initial state S_0 - point on line AB in plane xOy with the coordinates (x_1, y_1, 0) - into point K with the coordinates

$$x_1'=\varphi_1(x_1)\leqslant x_1,$$
$$y_1'=\psi_1(y_1)\leqslant y_1,$$
$$\xi_1=f_1(x_1)+g_1(y_1).$$

Then on the first component/link of second stage (2_1) occurs the enclosure of the part of the income and the redistribution of the resources between branches I and II. Point S is moved again to plane xOy into point M with the coordinates (x_2, y_2, 0), moreover

$$x_2+y_2=x_1'+y_1'+R_1(\xi_1).$$

Further again goes the expenditure of resources and the acquisition of income (component/link 2_2), then again redistribution, etc.

Our task - of deducing point S, which represents the state of system, on the plane

$$x+y+\xi=C$$

with the highest possible value of parameter C.

Let us sketch the diagram of the solution of problem by the method of dynamic programming.

Let us note first of all, that if is fixed/recorded issue (i-1) -th stage, then for the following (the i-th) is essential only the total sum of the redistributed resources

$$Z_{i-1}=x'_{i-1}+y'_{i-1}+R_{i-1}(\xi_{i-1}).$$

and therefore despite the fact that the state of system was represented as point in the three-dimensional space, we will vary the values only of one parameter Z_{i-1}.

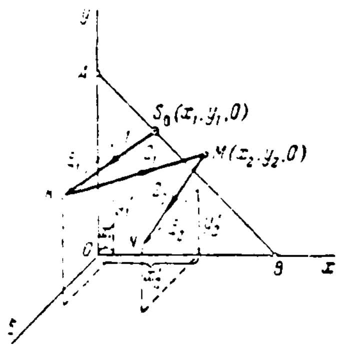


Fig. 11.2.

Page 97.

"control" on the i stage (just as in the previously tasks of distributing the resources/incomes examined) will consist of the selection of value x_i - quantity of resources, imbedded in branch I in the i stage. Prize W for entire process naturally is divided/marked off into m of the components/terms/addends:

$$W = w_1 + w_2 + \ldots + w_{m-1} + w_m. \quad (11.8)$$

where w_i with i=1, 2, ..., m-1 are the net income, not packed into the production:

$$w_i = \xi_i - R_i(\xi_i),$$

and at the m step/pitch - is entire net income from the m step/pitch plus the remainder/residue of the imbedded resources:

$$w_m = \xi_m + x'_m + y'_m.$$

Step by step optimization we will conduct according to the standard diagram.

1. We fix/record value Z_{m-1} (preserved rescurces plus packed part of income), which characterizes issue (m-1) -th of step/pitch. Conditional optimum control $x_m(Z_{m-1})$ on the m step/pitch will be located from the condition

$$W_m^*(Z_{m-1}) = \max_{0 < x_m \leq Z_{m-1}} \{w_m\} =$$
$$= \max_{0 \leq x_m \leq Z_{m-1}} \{f_m(x_m) + g_m(Z_{m-1} - x_m) +$$
$$+ \varphi_m(x_m) + \psi_m(Z_{m-1} - x_m)\}.$$

2. Let us fix issue of (m-2) -th of step Z_{m-2}. In order to find conditional optimum control on (m-1) -th step $x_{m-1}^*(Z_{m-2})$, necessary to maximize with the given one Z_{m-2} sum $W_{m-1,m}^+$ of the following values:

1) the remaining (not imbedded in the production) income at (m-1) -th step

$$w_{m-1} = f_{m-1}(x_{m-1}) + g_{m-1}(Z_{m-2} - x_{m-1}) -$$
$$- R_{m-1}(f_{m-1}(x_{m-1}) + g_{m-1}(Z_{m-2} - x_{m-1}));$$

2) prize at the latter/last step/pitch during the optimum control

$$W_m^*(Z_{m-1}) = W_m^*(\varphi_{m-1}(x_{m-1}) + \psi_{m-1}(Z_{m-2} - x_{m-1}) +$$
$$+ R_{m-1}(f_{m-1}(x_{m-1}) + g_{m-1}(Z_{m-2} - x_{m-1}))).$$

Thus, conditional optimum control on (m−1) -th the step/pitch is located as the value x_{m-1}, at which it is reached the maximum of value $W^{*}_{m-1,m}$:

$$W^{*}_{m-1,m}(Z_{m-2})=\max_{0\leq x_{m-1}\leq Z_{m-2}}\{W^{*}_{m-1,m}(Z_{m-2},\ x_{m-1})\}=$$
$$=\max_{0\leq x_{m-1}\leq Z_{m-2}}\{f_{m-1}(x_{m-1})+g_{m-1}(Z_{m-2}-x_{m-1})-$$
$$-R_{m-1}(f_{m-1}(x_{m-1})+g_{m-1}(Z_{m-2}-x_{m-1}))+$$
$$+W^{*}_{m}(\varphi_{m-1}(x_{m-1})+\psi_{m-1}(Z_{m-2}-x_{m-1})+$$
$$+R_{m-1}(f_{m-1}(x_{m-1})+g_{m-1}(Z_{m-2}-x_{m-1})))\}.$$

3. Conditional optimum control $x^{*}_{i}(Z_{i-1})$ on the i stage will be located from the relationship/ratio

$$W^{*}_{i,i+1,\ldots,m}(Z_{i-1})=\max_{0\leq x_i\leq Z_{i-1}}\{f_i(x_i)+g(Z_{i-1}-x_i)-$$
$$-R_i(f_i(x_i)+g_i(Z_{i-1}-x_i))+W^{*}_{i+1,\ldots,m}(\varphi_i(x_i)+$$
$$+\psi_i(Z_{i-1}-x_i)+R_i(f_i(x_i)+g_i(Z_{i-1}-x_i)))\}.$$

4. Optimum control x^{*}_{1} at first step/pitch and maximum value of prize W^{*} are found from condition

$$W^{*}=W^{*}_{1,2,\ldots,m}=\max_{0\leq x_1\leq Z_0}\{f_1(x_1)+g_1(Z_0-x_1)-$$
$$-R_1(f_1(x_1)+g_1(Z_0-x_1))+$$
$$+W^{*}_{2,\ldots,m}(\varphi_1(x_1)+\psi_1(Z_0-x_1)+R_1(f_1(x_1)+$$
$$+g_1(Z_0-x_1)))\}.$$

5. Issue of the first during optimum control

$$Z_1^* = \varphi_1(x_1^*) + \psi_1(Z_0 - x_1^*) + R_1(f_1(x_1^*) + g_1(Z_0 - x_1^*));$$

optimum control at second step/pitch:

$$x_2^* = x_2^*(Z_1^*);$$

issue of second step/pitch during optimum control:

$$Z_2^* = \varphi_2(x_2^*) + \psi_2(Z_1^* - x_2^*) - R_2(f_2(x_2^*) + g_2(Z_1^* - x_2^*)).$$

and so forth to latter/last step/pitch.

Page 99.

We recommend to reader to independently sketch the diagram of the solution of the following problems of distributing the resources/lifetimes.

1. To optimize distribution of resources according to two branches of production under following conditions: income is packed into production not completely, but partially ("function of enclosure" $R_i(\xi)$ $(i = 1, \ldots, m-1)$ are preset); is maximized total net income for all stages, without taking into account remaining resources.

2. To optimize distribution of resources according to two branches of production under following conditions: income is packed into production not completely, since its known fraction $\alpha_i(\xi)$ is

removed in the form of tax; remaining part is packed into production; it is maximized total quantity of resources (basic plus income) after m stage. There will not be any of these tasks under some conditions for that degenerated?

§ 12. Other varieties of the task of distributing the resources/lifetimes.

In this paragraph we will consider several tasks of the different regions of practice, which belong, actually, to the same category of "tasks for the distribution of the resources/lifetimes", but in which unusual setting immediately does not suggest about the familiar diagram. Calculating, that the reader already seized the principles of dynamic programming, we will allow ourselves with the solution of these problems on stepping back from standard notation, after preserving by constant/invariable only the diagram of the solution.

a. Task about weight distribution between steps/stages of space vehicle. One must plan multilevel space vehicle in the limits of the specific launching weight G. Cosmonaut's cabin has preset weight g_K. It is assumed that the rocket will have m of steps/stages.

Page 100.

Launching weight of rocket is composed of the weights of all steps/stages and cabin:

$$G = \sum_{i=1}^{m} G_i + g_K.$$

where G_i - weight of the i step/stage.

Each step/stage has some supply of combustible. After fuel depletion used-up stage it is discarded and it enters in the operation following.

Additional velocity Δv_i which acquires the rocket for the operating time of the engine of the i step/stage, depends both on the weight of step/stage itself G_i (being determining the fuel reserve) and on the weight of that cargo which it is necessary to carry:

$$\Delta v_i = f(G_i, P_i). \qquad (12.1)$$

where

$$P_i = G_{i+1} + G_{i+2} + \ldots + G_m + g_K \qquad (12.2)$$

- weight of the "passive" cargo, moved by the i stage of rocket.

It is necessary to find advantageous weight distribution $Q_0 = G - g_K$ between m stages of rocket, with which the velocity after the discharge/break of all steps/stages will be maximum.

Task is similar to one of the versions of the task of distributing the resources/lifetimes, namely - the task of the redundancy of the resources (see § 10, p. b). Actually/really, m of the stages of rocket it is possible to visualize as m of the stages of the process of acceleration. Before each stage we must solve: what part of being at our disposal weight, not spent, until now, we is spent to this stage, and what we reserve for the following. However, in comparison with the task of the redundancy of resources, examined into § 10, this task has certain special feature/peculiarity: function f, which is determining "income" from the stage of acceleration, it depends not on one argument - the "imbedded" resources, but from two - "imbedded" and "reserved". However, this does not vary the method of the solution and even it does not complicate it any substantially.

Let us designate G_i - weight, separated to the i-th step/stage ("control" in the i stage); $Q_i = Q_0 - (G_1 + G_2 + \dots + G_i)$ - weight, reserved to the remaining steps/stages. Value Q_i is analogous to the sum of resources Z_i that remains at our disposal after the i stage in the task about the redundancy of resources.

Page 101.

In the new designations formula (12.1) can be rewritten thus:

$$\Delta v_i = f(G_i, Q_i + g_K). \qquad (12.3)$$

Phase space, just as in the task about the redundancy of resources, can be assigned in the form of triangle AOB (Fig. 12.1). In each stage the trajectory reaches the axis of the ordinates (the "resources", isolated into the step/stage, completely are expended/consumed). Then S_0 lies/rests on line AB, point $S_{кон}$ - in the beginning of coordinates.

Let us begin, as always, from the latter/last stage. Any weight Q_{m-1}, which was preserved as a result of the previous stages, should be it goes without saying completely returned on m-th step/stage. Conditional optimum control at the m step/pitch will be

$$G_m^*(Q_{m-1}) = Q_{m-1}.$$

In this case will be acquired the conditional maximum velocity increment, which corresponds to given one Q_{m-1}:

$$\Delta V_m^*(Q_{m-1}) = f(Q_{m-1}, g_K).$$

We fix/record weight Q_{m-2}, which remained afterward (m-2) -th of stage. It is obvious,

$$Q_{m-1} = Q_{m-2} - G_{m-1}.$$

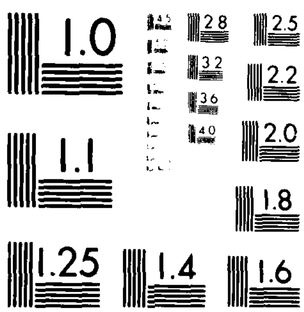
Conditional optimum control on (m - 1) - stage $G^{*}_{m-1}(Q_{m-2})$ will be located as rotating into the maximum the sum of two velocity increments: Δv_{m-1}, achieved/reached in (m-1) -th the stage with control G_{m-1}, and ΔV^{*}_{m} - maximum increase in the m stage:

$$\Delta V^{*}_{m-1,\,m}(Q_{m-2})= \max_{0<G_{m-1}\leq Q_{m-2}}\{f(G_{m-1},\,Q_{m-2}-G_{m-1}+g_K)+ +\Delta V^{*}_{m}(Q_{m-2}-G_{m-1})\}$$

and so on.

AD-A098 406 FOREIGN TECHNOLOGY DIV WRIGHT-PATTERSON AFB OH F/G 12/1
ELEMENTS OF DYNAMIC PROGRAMMING.(U)
FEB 81 Y S VENTTSEL'
UNCLASSIFIED FTD-ID(RS)T-1515-80 NL

3 OF 4
AD
A098406


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

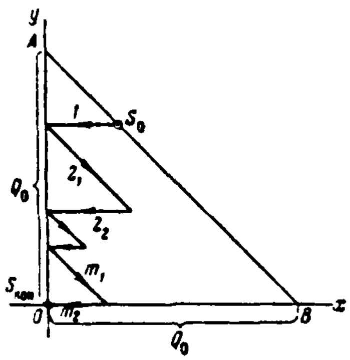


Fig. 12.1.

Page 102.

Conditional optimum control on the i step/pitch is found from the condition

$$\Delta V^*_{i,\, i+1,\, \ldots,\, m}(Q_{i-1}) = \max_{0<G_i<Q_{i-1}} \{f(G_i,\, Q_{i-1}-G_i+g_K) + \Delta V^*_{i+1,\, \ldots,\, m}(Q_{i-1}-G_i)\}.$$

After the optimization of the first step/pitch (selection of the weight of first stage G^*_1) the sequence of stages, as always, passes for a second time from the beginning toward the end; as a result is found the set of the optimum weights of the steps/stages:

$$G_1^*,\ G_2^*,\ \ldots,\ G_m^*;\quad \sum_{i=1}^{m} G_i^* = Q_0,$$

the imparting to the useful step/stage (cabin) maximum speed

$$\Delta V^* = \Delta V^*_{1,\, 2,\, \ldots,\, m}.$$

b. Distribution of weapons of destruction according to defended targets. In those tasks distributions of the resources/lifetimes which were encountered to us, until now, the resources, isolated in any stage, or gave income and due to this were expended (in full or in part), or they were reserved, they did not give income, but were not expended.

Here we will consider the peculiar task in which the resources are expended not only in that stage where they give "income", but also in those stages where they "income" do not yield, intensity of the expenditure of these resources depending on that, was how much imbedded in this stage of the directly functioning resources. Discussion deals with the distribution of resources/lifetimes with the "mutual support". As an example we will consider the task about the distribution of the resources of striking the defended targets.

Task is placed with following manner: is planned/glided the combat interaction by the specific weapons of destruction (for example, aircraft, rocket, winged missiles) on some defended targets (for example, ships, the anti-aircraft guns, etc.). Targets are distributed in depth in the depth of territory on several parallel borders of defense (Fig. 12.2).

Page 103.

Before to leave tc this border, weapons cf destruction pass zone the operaticns weapons of this border where they underge bombardment from the side of the latter. Weapons of each border can conduct fire/light not cnly according to the weapons cf destruction, which are guided directly for targets of this border, but alsc on those weapons of destructicn which pass through the zcne cf action, being directed to the more distant targets, arranged/located on the fcllowing borders.

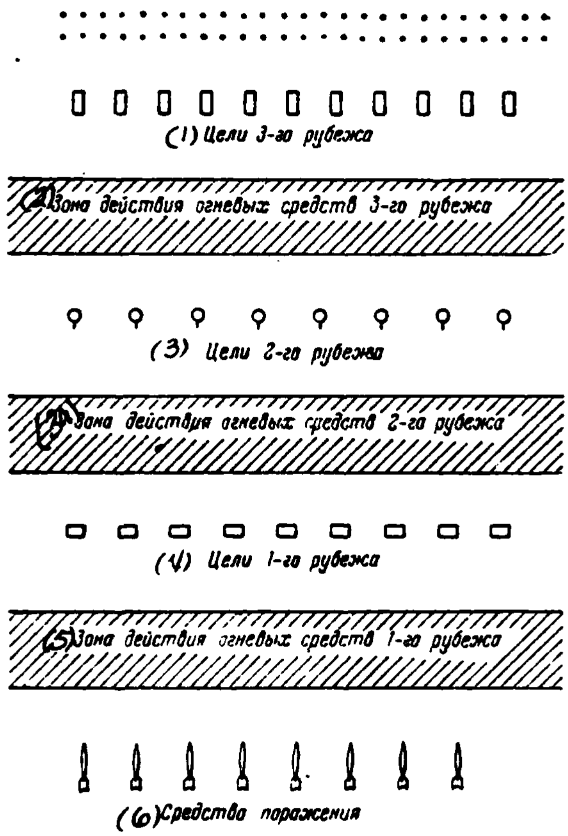


Fig. 12.2.

Key: (1). Targets of the 3rd border. (2). Zone of action weapons of 3rd border. (3). Targets of 2nd border. (3a). Zone of action weapons of 2nd border. (4). Targets of 1st border. (5). Zone of action weapons of 1st border. (6). Weapons of destruction.

Page 104.

The coating of weapons of destruction is planned/glided as follows: they are divided into the consecutive "waves"; the first wave is directed to the target of the 1st border, the second - on the target of the 2nd border, etc. The first wave passes through the zone of action weapons of the 1st border, it bears there known losses, after which the remaining weapons of destruction attack the targets of the 1st border, as a result of which some fraction/portion of these targets is surprised, and their weapons go out of order. Thus, after the coating of the first wave the 1st border of defense proves to be partially suppressed. Then enters in the operation the second wave; it moves through the zone of action of the partially suppressed weapons of the 1st border, loses the there certain part of its composition, then it enters into the zone of action weapons of the 2nd border, again loses there certain part of its composition; the remaining weapons of destruction attack the targets of the 2nd border, etc.

The task of planning the coating is posed as follows:

To distribute the available weapons of destruction on the waves so as to turn into the maximum average/mean number of the affected targets on all borders.

The posed problem by nature reminds of already familiar us the task of distributing the resources/lifetimes (" rescurces/lifetimes" are here weapons of destruction, by "income" - affected targets), but it differs from it in terms of two special features/peculiarities.

First, the weapons of destruction, isolated for the interaction on the targets of one or the other border, not only implement their primal problem (strike targets), but also they project/emerge as the "support" to the following waves, facilitating for them the overcoming the preliminary borders of defense.

In the second place, in contrast to all those it is previously examined, this task contains the element of randomness. Actually/really, an actual number of affected targets and failing weapons of destruction can prove to be the fact, etc. in the dependence on the random factors (for example, detection range, the accuracy of shooting, the failures of equipment, etc.).

Page 105.

The tasks of the dynamic programming, which contain the random factors (the so-called "stochastic" tasks) form special class and require the special approach (see § 15, 16). However, in this case we will not use this general/common/total approach, but solve task approximately with the help of the simplest method, frequently used in the similar cases: we will replace all figuring in the task random variables (number of affected targets on each border, number failing weapons of destruction) with their average/mean values (mathematical expectations). This method, which strongly simplifies task, usually gives comparatively small errors in the case when a number of the combat units (targets, weapons of destruction), which participate in the process, is sufficiently great [1].

FOOTNOTE [1]. An example of the task, decided not according to the "average/mean" characteristics, but with the real account to randomness, is given into § 16. ENDFOOTNOTE.

The solution of stated problem of distributing the weapons of destruction according to the defended targets simpler will consider based on specific example, after assigning the specific form of the figuring in it functional dependences.

Let be planned/glided coating n of aircraft on the air defense weapons (the anti-aircraft guns), arranged/located on m borders (Fig. 12.3).

In all on m borders there are by N of the instruments

$$N=\sum_{i=1}^{m} N_i. \qquad (12.4)$$

where N_i - number of instruments, arranged/located on the i border.

At our disposal there are by n of the aircraft from which must be formed with m of the waves:

$$n=\sum_{i=1}^{m} n_i. \qquad (12.5)$$

where n_i (i=1, 2, ..., m) - a number of aircraft, which form part of the i wave and which have the combat mission to influence on the instruments of the i border.

It is assumed that the waves are formed/shaped and is obtained the combat mission previously, and in the process of coating no longer they are reconstructed. Each wave flies before those following with certain prevention/advance on the time, so that up to the moment/torque of the approach of the following wave manages to already fulfill its combat mission.

Page 105.

Before emerging at the border of the location of instruments, each aircraft passes the zone of action of the instruments of this torder where it undergoes bombardment from the side of those instruments of this border which have the capability to shoot (i.e. they are found within reach and up to the given moment/torque they are not affected). To attack the instruments, arranged/located on this border, can only those aircraft, which happily passed the zone of the operation of the instruments of this border and all preceding.

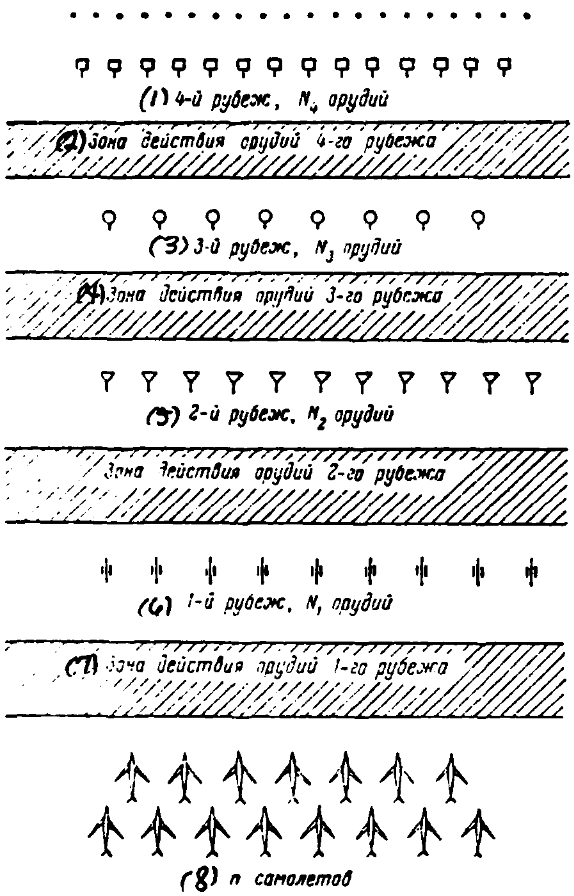


Fig. 12.3.

Key: (1). the 4th border, N_4 of instruments. (2). Zone of effect of instruments of 4th border. (3). 3rd border, N_3 of instruments. (4). Zone of action of instruments of 3rd border. (5). 2nd border, N_2 of instruments. (6). 1st border, N_1 of instruments. (7). Zone of action of instruments of 1st border. (8). p of aircraft.

Page 107.

Characteristics of the efficiency of the combat action of instruments on the aircraft and the aircraft on the instruments following.

1. Kill probability of one aircraft, which flies zone of action of instruments of i border, is expressed by formula

$$V_i = 1 - e^{-\alpha_i \tilde{N}_i}, \qquad (12.6)$$

where $\tilde{N}_i$ - average number of instruments, which were preserved by those nonafflicted on this border, α_i - coefficient, depending on efficiency of shooting of instruments at aircraft.

2. Average number of instruments of i border, beaten with wave aircraft, directed along targets of this border, is expressed by formula

$$Q_i = N_i \left[1 - e^{-\frac{\tilde{\nu}_i}{N_i} p_i}\right]. \qquad (12.7)$$

where N_i - number of instruments on i border, $\tilde{\nu}_i$ - average number of aircraft in i wave, which were preserved by those nonafflicted after passage of zones of action of instruments of this border and all previous, p_i - average/mean kill probability of one instrument cf border by its attacking aircraft.

It is necessary to assign the composition of waves, i.e., numbers n_1, n_2, ... n_m so as to become maximum an average number of affected targets on all borders:

$$W = \sum_{i=1}^{m} w_i.$$

where w_i - average number of affected targets of the i border.

In order to use the method of dynamic programming, it is necessary to, first of all, divide the planned/glide process into the steps/pitches (stages). This distribution can be made, generally speaking, by different methods; it is important only in the course of reasonings clear to visualize the determination of "step/pitch accepted" and not to be brought down from it to another.

We will divide process into the steps/pitches, on the basis of its following (it can be, sufficiently artificial) schematization.

Page 108.

Let us visualize that the zone of action of the instruments of the i border approaches certain number of aircraft Z_{i-1}, which happily surmounted all i - 1 previous borders of the defense (this number

Z_{i-1} we will assume/set by the equal to its average/mean value and allow/assume, thus, not only wholes, but also fractional "quantities of aircraft"). It is necessary this value - existing at our disposal of resource - to divide into two parts: x_i - the aircraft, which are guided for the damage/defeat of the instruments of the i border, and $y_i = Z_{i-1} - x_i$ - aircraft, "reserved" for the damage/defeat of the instruments of the subsequent borders. The first will meet nonweakened fire/light of the instruments of the i border, the second - with the fire/light, already weakened by previous interaction x_i of aircraft.

Upon this formulation of the problem we learn in it the already familiar signs/criteria of the task about the redundancy of resources.

Let us plan the overall diagram of its solution by the method of dynamic programming.

1. We fix/record result of (m-1) -th of step/pitch: zone of action of instruments of m border approached Z_{m-1} aircraft.

It is obvious, all these aircraft must be directed toward the damage/defeat of the instruments of the m border. Conditional optimum control at the m step/pitch will be

$$x_m^*(Z_{m-1}) = Z_{m-1}. \quad (12.8)$$

Let us determine the appropriate conditional maximum value of a number of affected instruments of m border $W^*_m(Z_{m-1})$. Since the m border yet did not undergo effect, on it were preserved all N_m of the instruments:

$$\tilde{N}_m = N_m. \qquad (12.9)$$

The kill probability of each of the chosen aircraft, according to formula (12.6), is equal to

$$V_m = 1 - e^{-\gamma_m N_m},$$

and an average number of aircraft which will happily cross the zone of the operation of the instruments of this border, it will be

$$\tilde{\nu}_m(Z_{m-1}) = Z_{m-1} \cdot (1 - V_m) = Z_{m-1} e^{-\gamma_m N_m}. \quad (12.10)$$

Page 109.

According to formula (12.7) these aircraft will strike the average number of instruments of the m border, equal to

$$W^*_m(Z_{m-1}) = N_m \left[1 - e^{-\frac{\tilde{\nu}_m}{N_m} p_m}\right]. \qquad (12.11)$$

where $\tilde{\nu}_m$, as shows formula (12.10), depends on Z_{m-1}.

Thus, on the m step/pitch are found conditional optimum control (12.8) and conditional maximum prize (12.11).

2. For planning/gliding (m-1)-th step/pitch we fix/record results (m-2) -th. let the zone of action of instruments (m-1) -th of border approach Z_{m-2} aircraft; from them it is necessary to isolate x_{m-1} on target (m-1) -th border, and the others to direct toward the m border through the zone of action of instruments of (m-1) -th border.

Conditional optimum control $x^{*}_{m-1}(Z_{m-2})$ will be located from the condition of maximum prize on two latter/last steps/pitches

$$W^{*}_{m-1,\,m}(Z_{m-2})= \\ = \max_{0 < x_{m-1} < Z_{m-2}} \{Q_{m-1}(x_{m-1}) + W^{*}_{m}(Z_{m-1})\}. \quad (12.12)$$

where $Q_{m-1}(x_{m-1})$ - average number of targets, affected on (m-1) -th border by those isolated for this x_{m-1} by aircraft; Z_{m-1} - average number of aircraft which will approach the zone of action of the instruments of the m border during this control (this value it depends both on the control at (m-1) -th step x_{m-1}, and from the number of aircraft $y_m = Z_{m-2} - x_{m-1}$, isolated into the flight/span of the zone of operation of (m-1) -th border).

According to formula (12.7) we have

$$Q_{m-1}(x_{m-1}) = N_{m-1}\left[1 - e^{-\frac{\tilde{x}_{m-1}}{N_{m-1}} r_{m-1}}\right]. \quad (12.13)$$

where

$$\tilde{x}_{m-1} = x_{m-1}(1 - V_{m-1}) = x_{m-1}e^{-\gamma_{m-1}N_{m-1}}. \quad (12.14)$$

Page 110.

Let us count an average number of nonafflicted aircraft of the second ("reserved") group, passing through the zone of action of instruments of (m-1) -th or border in order to be thrown from the m-th. Upon the input into the zone of their operation it was

$$Z_{m-2}-x_{m-1}.$$

An average number of instruments, affected on (m-1) -th border by aircraft, will be equally to $Q_{m-1}(x_{m-1})$. determined according to formula (12.13); consequently, on (m-1) -th the border will be preserved the average number of instruments, equal to

$$\tilde{N}_{m-1}=N_{m-1}-Q_{m-1}(x_{m-1}). \qquad (12.15)$$

These instruments by their fire/fight on the "flying" $Z_{m-2}-x_{m-1}$ aircraft decrease their number on the average to value

$$Z_{m-1}=(Z_{m-2}-x_{m-1})\cdot e^{-\epsilon_{m-1}\tilde{N}_{m-1}}. \qquad (12.16)$$

This value Z_{m-1}, which depends on Z_{m-2} and x_{m-1}, must be substituted into formula (12.12) and, varying control x_{m-1}, to find maximum conditional prize $W^{*}_{m-1,m}(Z_{m-2})$ and corresponding to it optimum conditional control $x^{*}_{m-1}(Z_{m-2})$.

In view of a comparative complexity of the figuring in the task functions hardly has the sense to attempt to seek maximum analytically; it is necessary to construct the series of the curves of the dependence of value

$$W^{+}_{m-1,m}=Q_{m-1}(x_{m-1})+W^{*}_{m}(Z_{m-1}). \qquad (12.17)$$

of that standing in the curly braces in right side (12.12), cf x_{n-1}. Each curve will correspond to that determined Z_{n-2} and on it will have to find point with the maximum ordinate. The abscissa of this point will be conditional optimum control at (m-1) -th step $x^*_{m-1}(Z_{m-2})$ and ordinate - corresponding to it conditional maximum income $W^*_{m-1,m}(Z_{m-2})$ at two latter/last steps/patches.

Will be further optimized (m-2) -th step etc.

Page 111.

3. General formulas for conditional maximum prize $W^*_{i,i+1,\ldots,m}(Z_{i-1})$ (and respectively conditional optimum control $x^*_i(Z_{i-1})$) take form

$$W^*_{i,i+1,\ldots,m}(Z_{i-1}) = \max_{0<x_i<Z_{i-1}} \{Q_i(x_i) + W^*_{i+1,\ldots,m}(Z_i)\}. \tag{12.18}$$

where

$$Q_i(x_i) = N_i\left[1 - e^{-\frac{\tilde{\gamma}_i}{N_i} n_i}\right]. \tag{12.19}$$

$$\tilde{\gamma}_i = x_i e^{-\gamma_i N_i}, \tag{12.20}$$

$$Z_i = (Z_{i-1} - x_i) e^{-\gamma_i \tilde{N}_i}, \tag{12.21}$$

$$\tilde{N}_i = N_i - Q_i(x_i). \tag{12.22}$$

4. According to general/common/total rule the process of

optimization continues right up to the first step/pitch, after which is sought optimum control at each step/pitch:

$$x_1^*, x_2^*, \dots, x_m^*.$$

However, the obtained numbers yet are not (with exception x_1^*) of the unknown optimum clusters of waves:

$$n_1^*, n_2^*, \dots, n_m^*.$$

since they are formed taking into account the losses of aircraft on all previous borders. In order, using x_i^* to find an initial number of aircraft in i wave n_i^* it is necessary to adjoin to x_i^* the average/mean losses of this wave δ_i^* on all previous borders.

Let us demonstrate the procedure of the optimization of control, after assigning the concrete/specific/actual numerical values of the parameters, which figure in the task.

Number of borders: m=4.

Number of aircraft: n=60.

Number of instruments on the borders: $N_1=10$; $N_2=12$; $N_3=15$; $N_4=10$ (in all $N=10+12+15+10=47$).

Kill probabilities of instrument by its one attacking aircraft:

$$p_1=0.4;\quad p_2=0.5;\quad p_3=0.4;\quad p_4=1.0.$$

Characteristics of the efficiency of the fire/light of air defense weapons on the aircraft:

$$\alpha_1 = 0.05; \quad \alpha_2 = 0.04; \quad \alpha_3 = 0.04, \quad \alpha_4 = 0.05.$$

Page 112.

To find optimum numbers of aircraft in the waves:

$$n_1^*, \quad n_2^*, \quad n_3^*, \quad n_4^*,$$

with which a number of affected instruments on all borders will be maximal.

The solution we will construct in stages.

1. Conditional optimization of fourth step/pitch. We are assigned by the series/row of the values of a number of aircraft Z_3, which approached the zone of action of the instruments of the 4th border, for example:

$$Z_3 = 10, 20, 30, 40, 50,$$

and let us compute for them an average number of struck instruments of the 4th border according to formula (12.11). The results of calculation let us design in the form of the graph/diagram of dependence $W^*_4(Z_3)$ (Fig. 12.4). Entering into this graph with any Z_3, we will be able to find the appropriate conditional maximum prize $W^*_4(Z_3)$; as far as control is concerned conditional optimum, then it is simply equal

$$x_4^*(Z_3) = Z_3.$$

2. Conditicnal optimization of third step/pitch. We are assigned by the series/row of the values of number Z_2 of the aircraft, which surmounted the previcus two borders:

$$Z_2 = 10,\ 15,\ 20,\ 25,\ 30,\ 40,$$

and for each of them let us count prize at twc latter/last steps/pitches: the third - during any ccntrcl and the fourth - with the optimum:

$$W_{3,4} = Q_3(x_3) + W_4^*(Z_3). \qquad (12.23)$$

where

$$Z_3 = (Z_2 - x_3)\, e^{-\gamma_3 \tilde{N}_3}; \qquad (12.24)$$

$$\tilde{N}_3 = N_3 - Q_3(x_3);$$

$$Q_3(x_3) = N_3\left[1 - e^{-\frac{\tilde{\nu}_3}{N_3} r_3}\right]; \qquad (12.25)$$

$$\tilde{\nu}_3 = x_3 e^{-\gamma_3 N_3}.$$

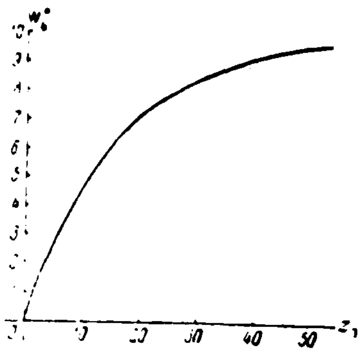


Fig. 12.4.

Page 113.

Value $Q_3(x_3)$, entering in (12.23), is counted according to formula (12.25), and $W^*_4(Z_3)$ is located through the graph/curve Fig. 12.4, for which it is necessary to enter into it with value of Z_3, undertaken from formula (12.24).

After producing calculations according to these formulas for the selected values of Z_2 and the series/row of values x_3, we construct the series of curves for function $w^+_{3,4}$ depending on x_3 (Fig. 12.5). For each of these curves we note point with the maximum ordinate. The abscissa of this point - conditional optimum control $x^*_3(Z_2)$, which corresponds to that Z_2, which includes the curve; ordinate - the appropriate conditional maximum prize $W^*_{34}(Z_2)$.

We construct on the graph of Fig. 12.6 (on the different scales) two curves: dependence $W^*_{34}(Z_2)$ and dependence $x^*_3(Z_2)$. The problem of the conditional optimization of the third step/pitch is solved.

3. Conditional optimization of second step/pitch. Procedure is completely analogous and shown in Fig. 12.7 and 12.8. First is constructed the series of curves $W^+_{Z_1 34}(x_2)$, which correspond to the different values of Z_1 (number of aircraft, which approached the zone of action of the instruments of the 2nd border):

$$Z_1 = 10,\ 20,\ 30,\ 40,\ 50,\ 60,$$

depending on control x_2 at the second step/pitch.

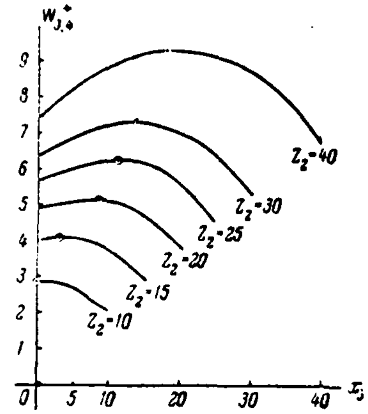


Fig. 12.5.

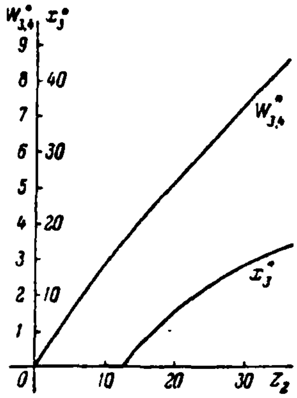


Fig. 12.6.

Pages 114-115.

Calculations are conducted according to the formula

$$W^{+}_{2,3,4} = Q_2(x_2) + W^{*}_{3,4}(Z_2). \qquad (12.26)$$

where

$$Q_2(x_2) = N_2\left[1 - e^{-\frac{\tilde{v}_2}{\tilde{N}_2} n_2}\right]. \qquad (12.27)$$

$$\tilde{v}_2 = x_2 e^{-\alpha_2 \nu_2}, \qquad (12.28)$$

and value $W^{*}_{3,4}(Z_2)$ is removed/taken from the graph of Fig. 12.6 with

$$Z_2 = (Z_1 - x_2)\, e^{-\alpha_2 \tilde{N}_2}, \qquad (12.29)$$

$$\tilde{N}_2 = N_2 - Q_2(x_2). \qquad (12.30)$$

For each of plotted curves (Fig. 12.7) again is located the maximum, and are constructed the dependences of its abscissa (conditional optimum control at the second step/pitch) and its ordinate (corresponding prize) on Z_1 (Fig. 12.8).

Graphing of Fig. 12.8 solved the task of the conditional optimization of the second step/pitch.

4. Optimization of first step/pitch. Value Z_0, with which we arrived at the optimization, are preset:

$$Z_0 = n = 80,$$

therefore we must construct only one curve dependence of $W_{1,2,3,4}$ on control x_1 at the first step/pitch (Fig. 12.9).

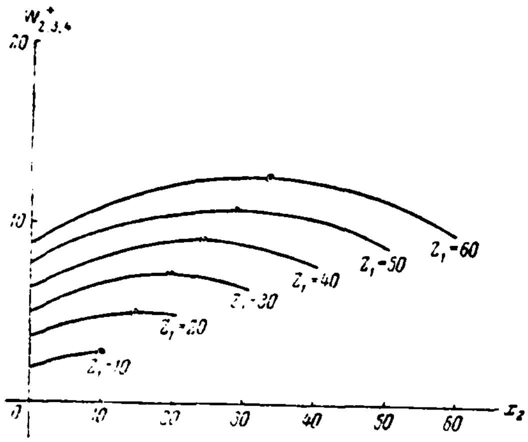


Fig. 12.7.

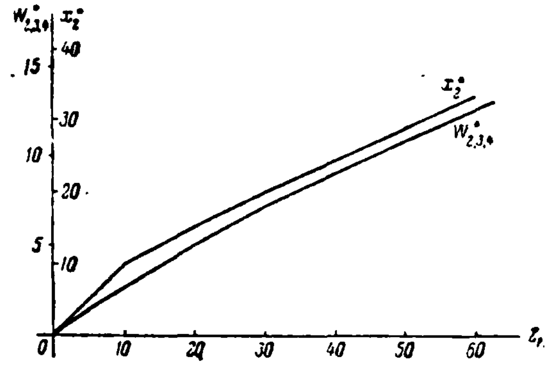


Fig. 12.8.

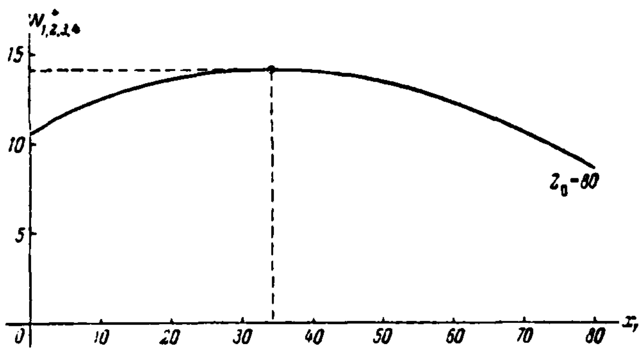


Fig. 12.9.

Page 116.

Calculations for the construction by this curve are conducted according to the formula

$$W^{+}_{1,2,3,4}(x_1) = Q_1(x_1) + W^{*}_{2,3,4}(Z_1). \qquad (12.31)$$

where

$$Q_1(x_1) = N_1\left[1 - e^{-\frac{\tilde{\nu}_1}{N_1}p_1}\right], \qquad (12.32)$$

$$\tilde{\nu}_1 = x_1 e^{-a_1 N_1}, \qquad (12.33)$$

but value $W^*_{2,3,4}(Z_1)$ is removed/taken from the graph of Fig. 12.8 with

$$Z_1 = (Z_0 - x_1)\, e^{-a_1 \tilde{N}_1}, \tag{12.34}$$
$$\tilde{N}_1 = N_1 - Q_1(x_1), \tag{12.35}$$
$$Z_0 = n = 80.$$

In the curve of Fig. 12.9 we note point with the maximum ordinate and thus we find a maximally possible prize (average number of affected instruments to all four borders)

$$W^* = W^*_{1,2,3,4} = 14.1$$

and optimum control on the first step/pitch

$$x_1^* \approx 34.$$

5. Optimization of entire process. We find optimum control step by step from the beginning to the end/lead. We isolated into the composition of the first wave (to the damage/defeat of the instruments of the 1st border) $x^*_1=34$ aircraft, and the others $y^*_1=80-34=46$ aircraft directed further.

The zone of action of the instruments of the 2nd border will approach (see formula (12.34)) the number of aircraft, equal to

$$Z_1^* = 46 e^{-a_1 \tilde{N}_1^*}, \tag{12.36}$$

where

$$\tilde{N}_1^* = N_1 - Q_1(x_1^*), \tag{12.37}$$

$$Q_1(x_1^*) = N_1 \left[1 - e^{-\frac{\tilde{x}_1^*}{N_1} p_1}\right], \tag{12.38}$$
$$\tilde{x}_1^* = x_1^* e^{-a_1 N_1}.$$

Producing calculations with a_1=0.05, N_1=10, we have

$$Q_1(x_1^*) = 5.6; \quad Z_1^* \approx 37.$$

i.e., on the 1st border it will be affected or the average of 5,6 instruments, and the zone of action of the instruments of the 2nd border it will approach on the average of 37 aircraft of 46 "that reserved".

Page 117.

With the obtained value of Z^*_1=37 let us enter into the graph of Fig. 12.8 and will find optimum control on the second step/pitch

$$x_2^* = 23,$$

i.e., from 37 preserved aircraft it is necessary to isolate 23 to the damage/defeat of the instruments of the 2nd border, and "to reserve"

$$y_2^* = 37 - 23 = 14.$$

We further find a number of aircraft which will approach the zone of action of the instruments of the 3rd border:

$$Z_2^* = 14e^{-\alpha_2 \tilde{N}_2^*}, \qquad (12.39)$$

where

$$\tilde{N}_2^* = N_2 - Q_2(x_2^*), \qquad (12.40)$$

$$Q_2(x_2^*) = N_2\left[1 - e^{-\frac{\tilde{v}_2^*}{N_2} p_2}\right], \qquad (12.41)$$

$$\tilde{v}_2^* = x_2^* e^{-\alpha_2 N_2}. \qquad (12.42)$$

Producing calculations according to these formulas, let us find an average number of instruments, beaten on the 2nd border:

$$Q_2(x_2^*) \approx 5.4.$$

and an average number of aircraft, which surmounted the first two borders:

$$Z_2^* = 10.7.$$

With value of Z^*_2=10.7 We enter into the graph of Fig. 12.6 and find optimum control on the third step/pitch

$$x_2^* = 0,$$

i.e., to the damage/defeat of the instruments of the 3rd border of aircraft to select completely not necessary ! This at first glance unexpected conclusion will be completely natural, if one considers that the shooting at the instruments of the 4th border under conditions of our task much more efficient than on the instruments of the 3rd border (p_3=0.4; p_4=1), and therefore has sense, in spite of the counteraction of the 3rd border, to reserve all 10.7 aircraft for the fourth wave.

From these 10.7 aircraft the zone of action of the instruments of the 4th border it will approach on the average

$$Z_3^* = Z_2^* \cdot e^{-a_3 V_3} = 5.9.$$

Page 118.

All of these 5,9 aircraft must be cast to the damage/defeat of the instruments of the 4th border. From them they will be preserved by those nonafflicted on the average

$$\tilde{v}_4^* = Z_3^* e^{-\gamma_4 N_4} = 3{,}6 \text{ самолета.}^{(1)}$$

Key: (1). aircraft.

and they will strike on the 4th border on the average

$$Q_4^* = N_4 \left[1 - e^{-\frac{\tilde{v}_4^*}{N_4} p_4}\right] = 3{,}1 \text{ орудия.}^{(1)}$$

Key: (1). instrument.

The process of optimization is completed: is found the optimum control:

$$x_1^* = 34; \quad x_2^* = 23; \quad x_3^* = 0; \quad x_4^* = 5{,}9.$$

It remains to pass from these values (quantity of aircraft, separated to this border from a number of those preserving to this border) to values $n_1^*, n_2^*, n_3^*, n_4^*$, included in the waves, formed/shaped in the beginning of coating. We have

$$n_1^* = x_1^* = 34,$$

i.e., in the first wave it is necessary to include/connect 34 aircraft of 80.

In what proportion it is necessary to divide the remaining 46 aircraft between the second and fourth waves? According to our calculations after the passage of the 1st border of 46 "reserved" aircraft it will remain 37, of them 23 must function on the instruments of the 2nd border. Since, according to condition, we must form/shape waves previously, but not on the borders of borders, then it is obvious, it is necessary to divide 46 the "reserved" aircraft between second and fourth waves in relationship/ratio 23:14, i.e., to include/connect in the second wave 27 aircraft, but in the fourth - remaining 19.

Thus,

$$n_1^* = 34; \quad n_2^* = 27; \quad n_3^* = 0; \quad n_4^* = 19.$$

During this optimum planning/gliding will be affected the maximally possible average number of targets, equal to

$$W^* = 14.1,$$

from bottom on 1st border 5,6, on 2nd 5,4, on 3rd not one and on 4th 3,1.

It does not represent the work to count also its own losses of aircraft with the execution of the combat mission. The part of these losses, namely losses in the "reserved" aircraft, we already computed in the course of computation, and for them it is necessary to supplement still less in those aircraft which, favorably after passing the preceded borders, are headed to the damage/defeat of the instruments of this border. An average quantity of these losses Π_i on the i border is computed from the formula

$$\Pi_i = x_i^*(1 - e^{-\gamma_i N_i}).$$

Page 119.

Applying this formula, and also taking into account the previously obtained losses in the reserved aircraft, we will obtain grand average losses:

on the 1st border

$$13.4 + 9.1 = 22.5;$$

on the 2nd border

$$8.8 + 3.3 = 12.1;$$

on the 3rd border

$$4.8;$$

on the 4th border

$$2.3.$$

Altogether of 80 aircraft with the fulfillment of operation on the suppression of air defense zone it will be lost in average/mean 42; by the price of such their own losses can be achieved/reached maximum "prize" - it is affected on the average of 14 instruments of the opponent.

It is obvious, such uncomforting of result does make it necessary to be planned about that, it is expedient to generally carry out operation on the suppression of the such well defended targets as in our example, with the help of such weapons of destruction as the examined by us aircraft? However, reasonings on this theme exceed the scope of the object/subject of dynamic programming, especially because initial numerical data, on which we constructed the solution, were selected from the purely systematic considerations and have nothing in common with the real ones.

Let us pause at one more question. Stated problem about the distribution of weapons of destruction we solve on the assumption that the distribution of aircraft into the waves and output of the combat mission to each wave is produced previously, and in the course of executing the operation its original plan does not vary.

Page 120.

In the principle the mission can be assigned otherwise: to assume that on the approaches to each border the actually preserved number of aircraft (which in the accuracy cannot be previously predicted) each time optimally is redistributed to two groups: one is directed to damage/defeat of the targets of this border, and another flies further. In the presence of the high speed control computer this optimum redistribution is quite possible. The exact solution of this task can be constructed with the general methods of the solution stochastic problems of the dynamic programming (see §§ 15, 16). However, in the first approximation, it is possible to use the following method.

In the beginning of combat (on overcomings of the 1st border) is solved the task of dynamic programming as this was shown above, and is located the optimum control on the first step/pitch x^*_1 which is realized. Then is discovered that the 2nd border it approached actually not Z^*_1 aircraft, but another number Z_1. But indeed with the solution of the problem of dynamic programming we for each Z_1 found the conditional optimum control $x^*_2(Z_1)$; we will use this dependence and let us use for that actually carrying out Z_1 this optimum control, etc.

Logically does arise the question: does have sense to study in the course of operation by this "redistribution", i.e., is great the prize in an average number of affected instruments, bought to the price of this complication of control? Response/answer to this question can be given, only if we construct the more exact "stochastic model" of combat operations (without the replacement of random values by their mathematical expectations), similar this is done into §§ 15, 16.

c. The same problem with another criterion. During the solution of the previous task (b) we assumed that the target of the operation, carried out by aircraft, only the damage/defeat of targets (instruments), the greater it will be defended affected these instruments, the better. Criterion W we have an average number of affected targets.

It is possible to consider another task when n of aircraft surmount air defense zone in order beyond its limits to fulfill some other, basic combat mission (for example, bombing on the industrial objects). In order to ensure a maximally successful fulfillment of this basic combat task, and is selected certain quantity of aircraft for the suppression of air defense weapons.

Page 121.

As the criterion during the evaluation of efficiency in the entire operation in this case it is wise to select an not average number of affected instruments, but an average number of aircraft Z_m, surmounted all m borders of PVO and ready to execution further combat operations:

$$W = Z_m.$$

This task has something general/common/total with the task about the distributions of resources/lifetimes, when is maximized not income, but the total quantity of resources (see § 11t), and just as that, proves to be the "degenerate" task of dynamic programming.

Actually/really in order to be convinced of this, let us sketch the diagram of the solution of problem by the method of dynamic programming. The criterion, which it is necessary to turn into the maximum, exists $W=Z_m$ - an average number of aircraft, happily surmounted all m borders.

1. We fix/record Z_{m-1}. Conditional optimum control at m step/pitch $x^*_m(Z_{m-1})$ no longer equal to Z_{m-1}, but it is found from the condition

$$W^*_m(Z_{m-1})=\max_{0<x_m<Z_{m-1}}\{Z_m(Z_{m-1},\ x_m)\},$$

where $Z_m(Z_{m-1},\ x_m)$ - average number of aircraft, which surmounted the m border with the preset number of aircraft Z_{m-1}, which enter to this border, and control x_m on this border.

Completely it is clear that Z_m is nondecreasing function Z_{m-1}, therefore, and W^*_m is also nondecreasing function Z_{m-1}.

2. We fix/record Z_{m-2}. Conditional optimum control on (m-1)-th step $x^*_{m-1}(Z_{m-2})$ will be located from the condition

$$W^*_{m-1,m}(Z_{m-2}) = \max_{0 < x_{m-1} < Z_{m-2}} \{W^*_m(Z_{m-1}(Z_{m-2}, x_{m-1}))\},$$

where $Z_{m-1}(Z_{m-2}, x_{m-1})$ - average number of aircraft, which surmounted the first m-1 borders during preset Z_{m-2} and fixed/recorded control x_{m-1}. It is clear that function $Z_{m-1}(Z_{m-2}, x_{m-1})$ is the nondecreasing function of argument Z_{m-2}; consequently, and $W^*_m(Z_{m-1}(Z_{m-2}, x_{m-1}))$ there is nondecreasing function Z_{m-2}, and this means, and $W^*_{m-1,m}(Z_{m-2})$ — the also nondecreasing function.

Page 122.

3. By completely analogous method let us ascertain that for any (i-th) step/pitch prize at all remaining steps/pitches to eat, nondecreasing function of number of aircraft Z_{i-1}, which approached this border.

Hence deducible is the conclusion: the posed problem is the "degenerate" task of dynamic programming. to plan is necessary each step/pitch separately, distributing the flown up to this border aircraft to two parts - "to the suppression of air defense weapons and "to the flight/span further" so as to become maximum an average number of aircraft, which surmounted this border (irrespectively of the others).

§ 13. Distribution of resources/lifetimes with the aftereffect.

Examined in § 12 tasks of distributing the rescurces/lifetimes were characterized by that special feature/peculiarity, that the means, isolated in one "branch", projected/emerged as the "support" for means, isclated in another, increasing the yielded by them income. This support was manifested so quickly after the enclosure of the corresponding means.

In practice can be met such situations, when activity of the "supporting" branch is manifested not right after enclosure in all means, but after some quantity of stages. In general an increase in the income of "basic" branch depends on that, how long passed from the moment/torque of the enclosure of means into the "supporting" branch.

The tasks of this type have, in comparison with those previously examined, certain feature: expedition from the "basic" branch in this stage depends not only on the state of system at the given moment/torque (where how much is imbedded), but also from the prehistory of the controlled process (where, when and how much were invested means). The presence of this "aftereffect" (effect of the past on the future) generally complicates the process of planning.

Earlier has already been discussed the fact that the presence of aftereffect it is possible to manage, including in the present state of system those parameters from the past, which essential for the future.

In this case increases the number of measurements of phase space and, therefore, sharply grows a quantity of versions of the states which must be sorted out in the process of optimization. However, fundamental difficulties this solution does not contain.

Let us consider a specific example or task for the distribution of means with the aftereffect.

Is planned/glided the work of enterprise with the initial supply of means Z_0 forward for period m of years. A quantity of means x, imbedded in the enterprise, is converted in a year (taking into account of income and expenditure of means) into some another quantity of means $F^{(0)}(x)$, which can be less, it is equal or more than initial x.

The means, available in the beginning of each year, we can at our discretion either completely pack into the production or partially expend/consume on the auxiliary actions, for example to the content of scientific laboratory, which, conducting research of production process, after certain time after its organization raises the profitableness of production, in consequence of which the function of a change in means $F^{(0)}(x)$ is substituted by another:

$$F^{(k)}(x) > F^{(0)}(x).$$

where k - quantity of years, during which the laboratory already existed.

Function $F^{(k)}(x)$ with any k we will consider nondecreasing.

On the content of laboratory, obviously, it is necessary to expend some means. Let us assume that these means - completely determined (they do not depend on our whim) and are equal to $\alpha(k-1)$ on the k year of the existence of laboratory (after it it worked already k-1 years). In this case $\alpha(0)$ indicates the original expenditures, required for the creation of laboratory and its content into course of the first year.

Let us agree to consider that if during some year of means to the laboratory they are not released, then it is preserved, achieved profitableness level of production is retained, but with the new enclosure of means laboratory functions and raises the profitableness of production in the manner that if interruption in the financing they was not.

Page 124.

"Control" of the distribution of means on each step/pitch (before beginning each new fiscal year) does consist of the solution of the question: tempering money to the laboratory or not to release? With this simplified formulation of task at each step/pitch is a selection only between two directions:

$U^{(0)}$ - not to release means,

$U^{(1)}$ - to release means.

It is necessary to find this control during m years, with which the total quantity of means, not imbedded in the laboratory (the net income plus the remaining part of the basic means), toward the end of the period will be maximal.

We will solve stated problem by the method of dynamic programming. It is first of all necessary to solve the question: by what parameters we will characterize the state of system afterward (i=1) -th of step/pitch (before beginning the i fiscal year)?

Is obvious, by one value Z_{i-1} - by a quantity of means, which are at our disposal afterward (i-1) -th of step/pitch, it is impossible to be bridged, since the income, which we will obtain at the i step/pitch, depends not only on what quantity of means we had available in the beginning of year and what control was used, but also from that, how many years, until now, worked the laboratory (with our assumptions nevertheless - worked it continuously or with the interruptions).

It is necessary to characterize the state of system after step/pitch to two parameters:

Z_{i-1} - a quantity of means (basic plus income), available afterward (i-1) -th step/pitch,

k_{i-1} - number of the years, during which functioned the laboratory before beginning the i step/pitch.

The state of system S_{i-1} afterward (i-1) -th step let us register in the form of vector with two components:

$$S_{i-1}=(Z_{i-1},\ k_{i-1}).$$

As the phase space let us consider on plane ZOk (Fig. 13.1) the series/row of straight/direct

$$0-0';\quad 1-1';\quad 2-2';\quad \ldots;\quad m-m',$$

parallel to axis abscissas; ordinates of these straight lines are equal to the integers: 0, 1, 2,... Along the axis of abscissas are plotted/deposited the distributed means Z, along the axis of ordinates - number of years of the existence of laboratory k.

Paga 125.

Initial state of syster - completely determination point S_0 on axis OZ with abscissa Z_0 (initial quantity of means). If at this step means to the laboratory are not released, point on the phase plane moves on the horizontal; if means to the laboratory are released, point is moved with this horizontal line to the following in order.

Region $\tilde{S}_{кон}$ of the final states of system is entire the phase space (set, straight lines 0-0', 1-1', etc.). Prize $W = Z_m$ is nothing else but the abscissa of latter/last point in the trajectory $S_{кон}$. The task of optimum planning/gliding is reduced to deduce the point, which represents the state of system, into final state $S_{кон}$ with the greatest abscissa (ordinate does not have a value).

Let us plan the diagram of the construction of optimum control by the method of dynamic programming. Let us represent gain $W = Z_m$ into form of sum m of the components/terms/addends:

$$W = w_1 + w_2 + \ldots + w_{m-1} + w_m,$$

from which all, except the latter, are equal to zero. We will construct the solution by the standard diagram.

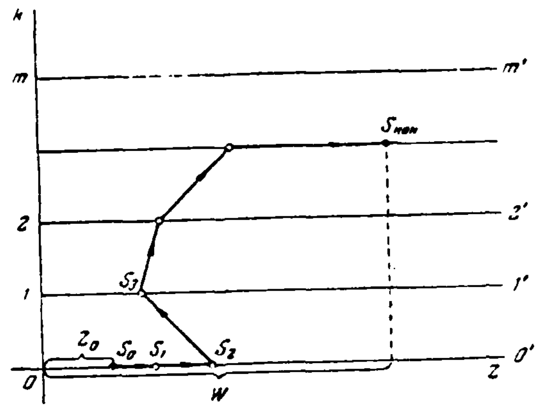


Fig. 13.1.

Page 126.

1. We fix/record result $(m-1)$ -th step/pitch, i.e., two numbers: quantity of means Z_{m-1} and number of years k_{m-1} during which already worked laboratory. Is obvious, $0 \leqslant k_{m-1} \leqslant m-1$. Let us consider and let us compare income at the m step/pitch which we will obtain during control $U^{(0)}$ (if we do not release money to the laboratory) and during control $U^{(1)}$ (if we release). In first the case in the production will be imbedded all means Z_{m-1}. and we will obtain the income

$$w_m(Z_{m-1}, k_{m-1}, U^{(0)}) = F^{(k_{m-1})}(Z_{m-1}). \quad (13.1)$$

In the second case in the production will be imbedded not all means, but only those which will remain after financing of laboratory, and we will obtain the income

$$w_m(Z_{m-1}, k_{m-1}, U^{(1)}) = F^{(k_{m-1})}(Z_{m-1} - a(k_{m-1})). \quad (13.2)$$

Since function $F^{(k_{m-1})}$ not decreasing, then it is obvious, of two incomes (13.1) and (13.2) the first is always more. Thus, optimum control at the latter/last step/pitch exists $U^{(0)}$ - not to release means to the laboratory, and this control does not depend on issue (m-1) -th of the step/pitch

$$U_m^* = U^{(0)},$$

but the corresponding maximum prize is equal to

$$W_m^*(Z_{m-1}, k_{m-1}) = F^{(k_{m-1})}(Z_{m-1}). \qquad (13.3)$$

2. We fix/record issue (m-2) -th step/pitch, i.e., vector

$$S_{m-2} = (Z_{m-2}, k_{m-2}).$$

During control $U^{(0)}$ the point in the phase space will pass on the horizontal into point S_{m-1} with coordinates

$$Z_{m-1} = F^{(k_{m-2})}(Z_{m-2}); \quad k_{m-1} = k_{m-2}.$$

Prize for the latter/last two steps/pitches (during the optimum control on the latter) will be

$$W_{m-1,m}(Z_{m-2}, k_{m-2}, U^{(0)}) = W_m^*(Z_{m-1}, k_{m-1}) =$$
$$= W_m^*(F^{(k_{m-2})}(Z_{m-2}), k_{m-2}), \quad (13.4)$$

or, using formula (13.3),

$$W_{m-1,m}(Z_{m-2}, k_{m-2}, U^{(0)}) = F^{(k_{m-2})}(F^{(k_{m-2})}(Z_{m-2})).$$
(13.5)

Page 127.

During control $U^{(1)}$ the point in the phase space will move with straight line $k_{m-2}-k'_{m-2}$ to the following in order straight line and will hit point S_{m-1} with the coordinates

$$Z_{m-1}=F^{(k_{m-2})}(Z_{m-2}-\alpha(k_{m-2}));$$
$$k_{m-1}=k_{m-2}+1. \qquad (13.6)$$

In this case $W^{+}_{m-1,m}$ it will be ascertained by the formula

$$W^{+}_{m-1,m}(Z_{m-2},\ k_{m-2},\ U^{(1)})=W^{*}_{m}(Z_{m-1},\ k_{m-1})=$$
$$=W^{*}_{m}(F^{(k_{m-2})}(Z_{m-2}-\alpha(k_{m-2})),\ k_{m-2}+1), \qquad (13.7)$$

or, taking into account (13.3),

$$W^{+}_{m-1,m}(Z_{m-2},\ k_{m-2},\ U^{(1)})=$$
$$=F^{(k_{m-2}+1)}(F^{(k_{m-2})}(Z_{m-2}-\alpha(k_{m-2}))). \qquad (13.8)$$

Optimum control ($U^{(0)}$ or $U^{(1)}$) will be located by comparison of two expressions (13.5) and (13.8) and with the selection of maximum of them

$$W^{*}_{m-1,m}(Z_{m-2},\ k_{m-2})=$$
$$=\max\left\{\begin{matrix}W^{+}_{m-1,m}(Z_{m-2},\ k_{m-2},\ U^{(0)}),\\ W^{+}_{m-1,m}(Z_{m-2},\ k_{m-2},\ U^{(1)})\end{matrix}\right\}.$$

3. For any (i-th) step/pitch optimum control ($U^{(0)}$ or $U^{(1)}$) will be located with comparison of two expressions

$$W^{+}_{i,i+1,\dots,m}(Z_{i-1},\ k_{i-1},\ U^{(0)})$$

and

$$W^{+}_{i,i+1,\dots,m}(Z_{i-1},\ k_{i-1},\ U^{(1)})$$

and with selection of maximum of them

$$W^{*}_{i,\,i+1,\,\dots,\,m}(Z_{i-1},\ k_{i-1}) = \\ = \max\begin{Bmatrix} W^{+}_{i,\,i+1,\,\dots,\,m}(Z_{i-1},\ k_{i-1},\ U^{(0)}), \\ W^{+}_{i,\,i+1,\,\dots,\,m}(Z_{i-1},\ k_{i-1},\ U^{(1)}) \end{Bmatrix}, \quad (13.9)$$

where

$$W^{+}_{i,\,i+1,\,\dots,\,m}(Z_{i-1},\ k_{i-1},\ U^{(0)}) = \\ = W^{*}_{i+1,\,\dots,\,m}(F^{(k_{i-1})}(Z_{i-1}),\ k_{i-1}), \quad (13.10)$$

$$W^{+}_{i,\,i+1,\,\dots,\,m}(Z_{i-1},\ k_{i-1},\ U^{(1)}) = \\ = W^{*}_{i+1,\,\dots,\,m}(F^{(k_{i-1})}(Z_{i-1} - x(k_{i-1})),\ k_{i-1}+1). \quad (13.11)$$

Page 128.

4. Optimization of first step/pitch is produced at fixed value of Z_0 and $k_0=0$:

$$W^{*} = W^{*}_{1,\,2,\,\dots,\,m} = \max\begin{Bmatrix} W^{+}_{1,\,2,\,\dots,\,m}(Z_0,\ 0,\ U^{(0)}), \\ W^{+}_{1,\,2,\,\dots,\,m}(Z_0,\ 0,\ U^{(1)}) \end{Bmatrix}. \quad (13.12)$$

where

$$W^{+}_{1,\,2,\,\dots,\,m}(Z_0,\ 0,\ U^{(0)}) = W^{*}_{2,\,\dots,\,m}(F^{(0)}(Z_0),\ 0), \quad (13.13)$$

$$W^{+}_{1,\,2,\,\dots,\,m}(Z_0,\ 0,\ U^{(1)}) = \\ = W^{*}_{2,\,\dots,\,m}(F^{(0)}(Z_0 - x(0)),\ 1). \quad (13.14)$$

Optimum control U_1^{*} at the first step/pitch will $U^{(0)}$ (drop means), if expression (13.13) is more than (13.14). If on the contrary, then optimum control will $U^{(1)}$ (release means [1]).

FOOTNOTE [1]. If expressions are equal, then both controls are equal.

ENDFOOTNOTE.

5. Further is located result of first step/pitch during optimum control: (Z_1^*, k_1^*); then optimum control at second step/pitch U_2^* and so on, up to latter/last step/pitch.

Let us demonstrate the solution of problem based on numerical example.

Let us assume m=4 and let us assign function $F^{(k)}(x)$ and $\alpha(k)$ with $0 \leq k \leq 3$:

$$
\begin{aligned}
&F^{(0)}(x) = 1.5x; && \alpha(0) = 1; \\
&F^{(1)}(x) = 1.6x; && \alpha(1) = 0.5; \\
&F^{(2)}(x) = 2x; && \alpha(2) = 0.4; \\
&F^{(3)}(x) = 3x; && \alpha(3) = 0.3.
\end{aligned}
$$

1. At latter/last step/pitch, as it is already explained,

$$U_m^* = U_4^* = U^{(0)},$$

i.e. means to laboratory to release not necessary. During this optimum control the prize at the fourth step/pitch for different k will be equal to

$$
\begin{aligned}
&W_4^*(Z_3, 0) = 1.5Z_3; \\
&W_4^*(Z_3, 1) = 1.6Z_3; \\
&W_4^*(Z_3, 2) = 2Z_3; \\
&W_4^*(Z_3, 3) = 3Z_3.
\end{aligned}
$$

Page 129.

2. We optimize third step/pitch. We have

$$W_{3,4}^{+}(Z_2,\ 0,\ U^{(0)}) = W_4^{*}(1.5Z_2,\ 0) = 2.25Z_2;$$
$$W_{3,4}^{+}(Z_2,\ 0,\ U^{(1)}) = W_4^{*}(1.5(Z_2-1),\ 1) =$$
$$= 1.6(1.5Z_2 - 1.5) = 2.4Z_2 - 2.4.$$

Of these two expressions the first large of the second with $Z_2 < 16$; when $Z_2 > 16$ - vice versa. Therefore

$$W_{3,4}^{*}(Z_2,\ 0) = \begin{cases} 2.25Z_2 & (1)\ \text{при}\ \ Z_2 \leqslant 16, \\ 2.4Z_2 - 2.4 & (1)\ \text{при}\ \ Z_2 \geqslant 16 \end{cases} \quad (13.15)$$

Key: (1). with.

and respectively

$$U_3^{*}(Z_2,\ 0) = \begin{cases} U^{(0)} & (1)\ \text{при}\ \ Z_2 \leqslant 16, \\ U^{(1)} & (1)\ \text{при}\ \ Z_2 \geqslant 16. \end{cases} \quad (13.16)$$

Key: (1). with.

Further we have

$$W_{3,4}^{+}(Z_2,\ 1,\ U^{(0)}) = W_4^{*}(1.6Z_2,\ 1) = 1.6 \cdot 1.6Z_2 = 2.56Z_2;$$
$$W_{3,4}^{+}(Z_2,\ 1,\ U^{(1)}) = W_4^{*}(1.6(Z_2 - 0.5);\ 2) =$$
$$= 2(1.6Z_2 - 0.8) = 3.2Z_2 - 1.6.$$

Of these two expressions the first will be more than the second with $Z_2 > 2.5$. Therefore

$$W_{3,4}^{*}(Z_2,\ 1) = \begin{cases} 2.56Z_2 & (1)\ \text{при}\ \ Z_2 \leqslant 2.5, \\ 3.2Z_2 - 1.6 & (1)\ \text{при}\ \ Z_2 \geqslant 2.5 \end{cases} \quad (13.17)$$

Key: (1). with.

and respectively

$$U_3^*(Z_2, 1) = \begin{cases} U^{(0)} \text{ (1) при } & Z_2 < 2.5, \\ U^{(1)} \text{ (1) при } & Z_2 \geqslant 2.5. \end{cases} \quad (13.18)$$

Key: (1). with.

Further,

$$W_{3,4}^*(Z_2, 2, U^{(0)}) = W_1^*(2Z_2, 2) = 4Z_2,$$
$$W_{3,4}^*(Z_2, 2, U^{(1)}) = W_1^*(2(Z_2 - 0.4), 3) = 6Z_2 - 2.4.$$

Of these two expressions the first is more with $Z_2>1.2$, the second - with $Z_2>1.2$; consequently,

$$W_{3,4}^*(Z_2, 2) = \begin{cases} 4Z_2 & \text{(1) при } Z_2 < 1.2, \\ 6Z_2 - 2.4 & \text{(1) при } Z_2 \geqslant 1.2 \end{cases} \quad (13.19)$$

Key: (1). with.

and respectively

$$U_3^*(Z_2, 2) = \begin{cases} U^{(0)} \text{ (1) при } & Z_2 < 1.2, \\ U^{(1)} \text{ (1) при } & Z_2 \geqslant 1.2. \end{cases} \quad (13.20)$$

Key: (1). with.

Page 130.

To assume/set $k_2=3$ no longer necessary, since for two first steps/pitches system cannot arrive to straight/direct 3-3'.

3. We optimize second step/pitch. We have

$$W_{2,3,4}^*(Z_1, 0, U^{(0)}) = W_{3,4}^*(1.5Z_1, 0),$$

or, using formula (13.5),

$$W^*_{2,3,4}(Z_1, 0, U^{(0)}) = \begin{cases} 3.375Z_1 & \text{(1) при } 1.5Z_1 \leqslant 16, \\ & \text{т. е. (1) при } Z_1 \leqslant 10\frac{2}{3} \approx 10.67, \\ 3.6Z_1 - 2.4 & \text{(1) при } Z_1 \geqslant 10\frac{2}{3} \approx 10.67. \end{cases} \quad (13.21)$$

Key: (1). with.

It is analogous

$$W^*_{2,3,4}(Z_1, 0, U^{(1)}) = W^*_{3,4}(1.5(Z_1 - 1), 1).$$

or, using formula (13.7),

$$W^*_{2,3,4}(Z_1, 0, U^{(1)}) = \begin{cases} 3.84Z_1 - 3.84 & \text{(1) при } 1.5(Z_1 - 1) \leqslant 2.5, \\ & \text{т. е. (1) при } Z_1 \leqslant 2\frac{2}{3} \approx 2.67, \\ 4.8Z_1 - 6.4 & \text{(1) при } Z_1 \geqslant 2\frac{2}{3} \approx 2.67. \end{cases} \quad (13.22)$$

Key: (1). with.

In order to solve a question, which of expressions (13.21) or (13.22) is more, let us construct the graphs of the corresponding functions (Fig. 13.2).

Each of the curves will be broken line, comprised of two straight lines. The maximum of these two functions is shown in Fig. 13.2 by heavy line. Letters $U^{(0)}$ and $U^{(1)}$ marked the curves, corresponding to the corresponding controls. Broken line, which represents the maximum of two functions (13.11) and (13.12), current consists of two segments; break point has an abscissa, equal to 4.43.

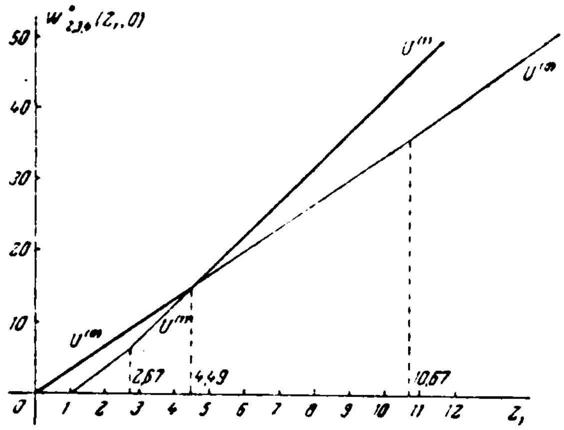


Fig. 13.2.

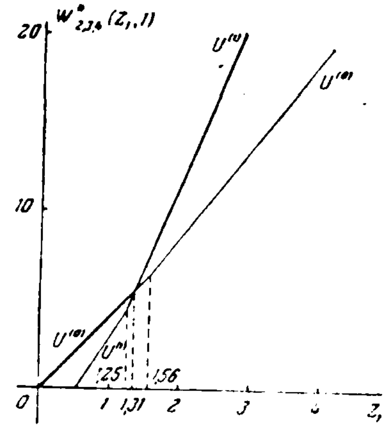


Fig. 13.3.

Page 132.

The equation of this curve will be

$$W^*_{2,3,4}(Z_1, 0) = \begin{cases} 3.375Z_1 & {}^{(1)}\text{при } Z_1 \leq 4.49, \\ 4.8Z_1 - 6.4 & {}^{(1)}\text{при } Z_1 > 4.49 \end{cases} \quad (13.23)$$

Key: (1). with.

and, therefore,

$$U^*_{2,3,4}(Z_1, 0) = \begin{cases} U^{(0)} & {}^{(1)}\text{при } Z_1 \leq 4.49; \\ U^{(1)} & {}^{(1)}\text{при } Z_1 > 4.49 \end{cases} \quad (13.24)$$

Key: (1). with.

We further find

$$W^*_{2,3,4}(Z_1, 1, U^{(0)}) = W^*_{3,4}(1.6Z_1, 1),$$

or, using formula (13.7),

$$W^*_{2,3,4}(Z_1, 1, U^{(0)}) =$$

$$= \begin{cases} 4.10Z_1 & {}^{(1)}\text{при } 1.6Z_1 \leq 2.5, \\ & {}^{(2)}\text{т. е. } {}^{(1)}\text{при } Z_1 \leq 1.56, \\ 5.12Z_1 - 1.6 & {}^{(1)}\text{при } Z_1 > 1.56. \end{cases} \quad (13.25)$$

Key: (1). with.

Further, $W^*_{2,3,4}(Z_1, 1, U^{(1)}) = W^*_{3,4}(1.6(Z_1 - 0.5), 2)$,
or, using formula (13.9),

$$W^*_{2,3,4}(Z_1, 1, U^{(1)}) = \begin{cases} 6.4Z_1 - 3.2 & \text{при}^{(1)} \quad 1.6Z_1 - 0.8 < 1.2, \\ & \text{т. е. при}^{(1)} \quad Z_1 \leqslant 1.25, \\ 9.6Z_1 - 7.2 & \text{при}^{(1)} \quad Z_1 \geqslant 1.25. \end{cases} \quad (13.26)$$

Key: (1). with.

Expressions (13.25) and (13.26) are represented graphically in Fig. 13.3. The maximum of two functions (13.25) and (13.26) is shown in Fig. 13.3 of fatty/greasy broken line of lines whose equation

$$W^*_{2,3,4}(Z_1, 1) = \begin{cases} 4.10Z_1 & \text{при}^{(1)} \quad Z_1 \leqslant 1.31, \\ 9.6Z_1 - 7.2 & \text{при}^{(1)} \quad Z_1 \geqslant 1.31; \end{cases} \quad (13.27)$$

Key: (1). with.

hence

$$U^*_2(Z_1, 1) = \begin{cases} U^{(0)} & \text{при}^{(1)} \quad Z_1 \leqslant 1.31; \\ U^{(1)} & \text{при}^{(1)} \quad Z_1 \geqslant 1.31. \end{cases} \quad (13.28)$$

Key: (1). with.

To assume/set $k_1=2$ no longer necessary, since for one first step/pitch system cannot arrive to straight/direct 2-2'.

4. We optimize first stop/pitch.

Page 133

We have

$$W^{+}_{1,2,3,4}(Z_0,\ 0,\ U^{(0)}) = W^{*}_{2,3,4}(1.5Z_0,\ 0),$$

or, using formula (13.13),

$$W^{+}_{1,2,3,4}(Z_0,\ 0,\ U^{(0)}) = \begin{cases} 5.06Z_0 & \text{(1) при } 1.5Z_0 \leqslant 4.49, \\ & \text{(2) т. е. (1) при } Z_0 \leqslant 2.93, \\ 7.27Z_0 - 6.4 & \text{(1) при } Z_0 \geqslant 2.93. \end{cases} \quad (13.29)$$

Key: (1). with.

It is analogous

$$W^{+}_{1,2,3,4}(Z_0,\ 0,\ U^{(1)}) = W^{*}_{2,3,4}(1.5(Z_0 - 1),\ 1),$$

or, accordingly, to formula (13.17),

$$W^{+}_{1,2,3,4}(Z_0,\ 0,\ U^{(1)}) = \begin{cases} 6.14Z_0 - 6.14 & \text{(1) при } 1.5Z_0 - 1.5 \leqslant 1.31, \\ & \text{(2) т. е. (1) при } Z_0 \leqslant 1.87, \\ 14.4Z_0 - 2.16 & \text{(1) при } Z_0 \geqslant 1.87. \end{cases} \quad (13.30)$$

Key: (1). with.

Expressions (13.29) and (13.30) are represented graphically in Fig. 13.4. Fatty/greasy of lines is shown the maximum of these two functions; the equation of this line

$$W^{*}(Z_0) = W^{*}_{1,2,3,4}(Z_0) = \begin{cases} 5.06Z_0 & \text{(1) при } Z_0 \leqslant 2.31, \\ 14.4Z_0 - 21.6 & \text{(1) при } Z_0 \geqslant 2.31. \end{cases} \quad (13.31)$$

Key: (1). with.

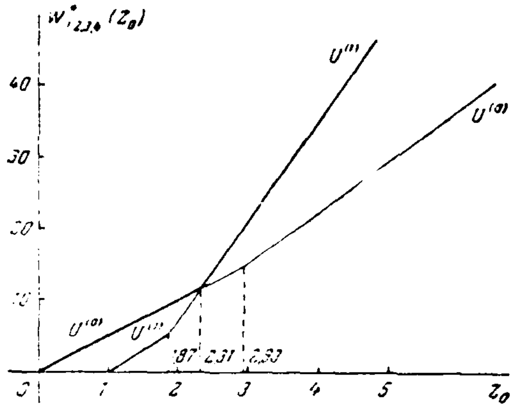


Fig. 13.4.

Page 134.

Hence optimum control at the first step/pitch will be

$$U_1^*(Z_0)=\begin{cases} U^{(0)} \text{ (1) при } & Z_0 \leqslant 2.31, \\ U^{(1)} \text{ (1) при } & Z_0 \geqslant 2.31. \end{cases} \quad (13.32)$$

Key: (1). with.

Thus, the conditional optimization of each step/pitch is carried out.

5. let us find now optimum control in all four years:

$$U^* = (U_1^*, U_2^*, U_3^*, U_4^*).$$

Let us consider two cases:

a) $Z_0<2.31$;

b) $Z_0>2.31$.

In the case of a) optimum control at the first step/pitch will be $U_1^*=U^{(0)}$, on the first year of means to the laboratory to release not necessary. To end of the first year we will have number of the means

$$Z_1^*=1.5Z_0<3.47;$$

in this case $k_1^*=0$, т. е. $S_1^*=(1.5Z_0,\ 0)$.

In order to find optimum control on second step/pitch U_2^*, let us turn to the graph of Fig. 13.2. Since $Z_1^*<3.47<4.49$, the optimum control again is $U^{(0)}$:

$$U_2^*=U^{(0)}.$$

i.e. on the second year of means to the laboratory to release not necessary.

We find the result of the second step/pitch during the optimum control:

$$Z_2^*=1.5Z_1^*=2.25Z_0<5.20;\quad k_2^*=0;$$
$$S_2^*=(2.25Z_0,\ 0).$$

From formula (13.6) follows that also at the third step/pitch the optimum control exists $U^{(0)}$ (not to release means), since

$$Z_2^* < 5.20 < 16.$$

Control at the latter/last step/pitch is standard: to release means to the laboratory is not necessary.

Page 135.

In this case the maximum prize (see formula (13.21)) will be equal to

$$W^* = 5.06 Z_0.$$

Thus, we ascertained that with an initial quantity of means Z_0<2.31 optimally control exists

$$U^* = (U^{(0)}, U^{(0)}, U^{(0)}, U^{(0)}),$$

i.e. laboratory generally started must not be.

Analogously we are convinced, that with Z_0>2.31 the optimum control will be

$$U^* = (U^{(1)}, U^{(1)}, U^{(1)}, U^{(0)}),$$

i.e. laboratory one should to bring immediately, and to last year preserve. In this case maximum income will be equal to

$$W^* = 14.4 Z_0 - 21.6.$$

Fig. 13.5 shows two optimum trajectories in the phase space for two concrete/specific/actual values of Z_0:

$$Z_0 = 2.2 < 2.31 \quad и \quad Z_0 = 2.4 > 2.31.$$

The first trajectory corresponds to the case when means to the laboratory are not released; the second - to case when laboratory is financed in any stage, except the latter itself.

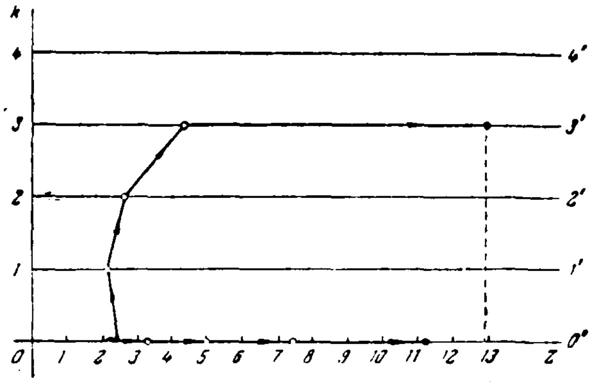


Fig. 13.5.

Page 136.

Let us note that in our example the optimum control, depending on the initial supply of means z_0, is constructed according to the type "all or nothing" y either with a sufficient supply of means it is necessary always (except last year) to hold laboratory, or, if the supply of means is insufficient, to in no way start it. Here not under such initial conditions there can be advantageous at first to accumulate means, and to then put them into the laboratory.

This is connected with the fact that we considered too small a number of steps/pitches (m=4). It is possible to propose to reader this exercise: to supplement our initial data, i.e., to assign function $F^{(4)}(x)$, $F^{(5)}(x)$, ... and expenditures/consumptions a(4), a(5),

..., and to attempt to obtain such solution during which optimum control will be "mixed": at the initial steps/pitches the means to the laboratory are not released, and then this becomes advantageous.

§ 14. Tasks of dynamic programming with the nonadditive criterion.

All tasks of the dynamic programming which we examined, until now, they belonged to the class of the "additive" tasks, i.e., such, in which it is maximized (it is minimized) the criterion of the form

$$W = \sum_{i=1}^{m} w_i, \qquad (14.1)$$

where w_i - the "prize", acquired in the i stage.

Generally the method of dynamic programming can be used also to the "nonadditive" tasks, in which the criterion is not represented in the form (14.1). Some of them are reduced to by the additive simple conversion of criterion.

Here we will consider the simplest form of the tasks, which are led to the the additive, namely task with the multiplicative criterion.

"Multiplicative" we will call criterion ("prize") W, if it can be represented in the form of the product of the "prizes", reached in the single stages:

$$W = w_1 w_2 \ldots w_m = \prod_{i=1}^{m} w_i. \qquad (14.2)$$

Page 137.

It is obvious, any multiplicative criterion W of form (14.2) can be artificially converted to the additive, if we take the logarithm expression W:

$$\lg W = \sum_{i=1}^{m} \lg w_i \qquad (14.3)$$

and to designate

$$V = \lg W; \quad v_i = \lg w_i. \qquad (14.4)$$

We will obtain the new criterion

$$V = \sum_{i=1}^{m} v_i, \qquad (14.5)$$

possessing the property of additivity and turning into the maximum (minimum) simultaneously with W.

Let us consider two examples of tasks with the multiplicative criterion.

a. Distribution n of projectiles according to m targets which must be struck together. Let there be at our disposal by n of the projectiles by which we wish to strike m of the targets:

$$\text{Ц}_1, \text{Ц}_2, \ldots, \text{Ц}_m \quad (m < n).$$

the combat mission lies in the fact that to strike all targets without the excepticn/eliminaticn. It is necessary so tc distribute n cf projectiles cn m targets so that the prctatility of the ccmbined damage/defeat of all targets W would reach maximum.

We will consider the distribution of projetiles as m- the stage operation, in each stage of which occurs the extraction of certain number of projectiles tc the specific target. Let us designate k_i a number of projectiles, isolated into the i-th target (i=1, 2, ..., m). Control u will consist of the selection cf the numbers $k_1, k_2, \ldots, k_m$:

$$U = (k_1, k_2, \ldots, k_m). \qquad (14.6)$$

and it is necessary tc select it optimally.

Let us assume that in the process of the bcmbardment of each target are expended/consumed all chosen intc it projectiles and the redistribution of means is not produced. Furthermcre, let us assume, that the discrete targets are surprised by the chosen on them projectiles independently of each other.

Page 138.

Then kill probability all m of targets is equal to the product of the kill probabilities of the discrete targets:

$$W = \prod_{i=1}^{n} P_i(k_i). \qquad (14.7)$$

where $P_i(k_i)$ - kill prctatility of the i target by isclated on carried by projectiles.

For the solution of posed problem it is required to assign the function

$$P_1(k),\ P_2(k),\ \ldots,\ P_m(k), \qquad (14.8)$$

characterizing the vulnerability of targets and expressing kill probability by each of them depending on number k of chosen into it projectiles.

If all these functions are identical:

$$P_1(k) = P_2(k) = \ldots = P_m(k) = P(k),$$

i.e. the vulnerability of all targets some and the same, then task becomes trivial and is reduced to distribute projectiles on the targets about the possibility evenly. But if targets are dissimilar by the vulnerability, obviously, the quantity of projectiles, separated on the basis of the less vulnerable targets, must be relatively more.

Before slinging the solution by the method of dynamic programming, let us note some of its properties.

Each of the functions $P_i(k)$, becomes zero with k=0, i.e.,

$$P_i(0) = 0 \text{ each } i = 1, 2, \ldots, m.$$

therefore, if we do not fire, at least one of the targets, criterion W will become zero. Hence follows the condition

$$k_i \geqslant 1 \quad (i = 1, 2, \ldots, m).$$

i.e. on each target it is necessary to isolate at least one projectile. Furthermore, selecting projectiles on the basis of the i target, we must remember about the fact that to those remaining m-i targets it is necessary without fail to reserve at least on one projectile; therefore each of the values k_i is limited not only from below, but also on top:

$$1 \leqslant k_i \leqslant n - m + 1. \qquad (14.9)$$

Page 139.

In order to use the method of dynamic programming, let us pass from multiplicative criterion (14.7) to the additive, after taking the logarithm it with any basis/base (for example, natural e):

$$\ln W = \sum_{i=1}^{m} \ln P_i(k_i). \qquad (14.10)$$

In order not to deal concerning negative numbers, let us designate

$$|\ln W| = V, \quad |\ln P_i(k_i)| = v_i(k_i) \qquad (14.11)$$

and we will obtain the raw - additive - criterion

$$V = \sum_{i=1}^{m} v_i(k_i). \qquad (14.12)$$

Since we changed the sign of criterion, then value V it is necessary to no longer to maximize, but to minimize.

Thus, tasks it is reduced to the following: to find control U^*, which rotates into the minimum value (14.12).

The obtained task calls to mind that examined into § 10 task of distributing the resources/lifetimes with the redundancy. Actually/really, available n of projectiles can be considered as the initial means which into each stage of operation are divided into two parts: separated and reserved, moreover the separating means are expended/consumed to the end/lead. The special feature/peculiarity of this task in the fact that, in the first place, number of the means, separated in each stage, can have only integer values, limited by condition (14.); furthermore, the "income" of V it is not maximized, but it is minimized.

In the tasks of distributing the resources/lifetimes we, for the uniformity, each time as the phase space examined triangle AOE on plane xOy. Here it would be possible to do so, but we will prefer another, more demonstrative method of the image of the process during which directly evidently not only control, but also the obtained "prize".

Let us consider as phase space the set of the straight/direct, parallel to axis abscissas: 0-0', 1-1',..., $n-n'$ with the integral ordinates (Fig. 14.1). Point S, which represents the state of system, will be in the process of the consumption of projectiles moved with one of these straight lines to another. If into this target is isolated only one projectile, then point S will be moved to the adjacent straight line; if two - are jumped through the straight line, etc.

Page 140.

Thus, along the axis of ordinates will be plotted/deposited the spent to this stage number of projectiles $N_{(n)}$. Along the axis of abscissas we will plot/deposit the accumulated for several stages "prize" $\sum v_i$.

Obviously, the initial state of system S_0 completely to determination coincides since the origin of the coordinates; region $\tilde{S}_{\kappa o \mu}$ of the final states of system is the straight line n-n' (in Fig. 14.1 n=9, m=5). It is necessary to find such control U (this trajectory in the phase space), during which abscissa $V=\sum_{i=1}^{m} v_i$ of end point S_{κ} will be smallest.

Solve to the end/lead stated problem of dynamic programming, being assigned by the specific values of n and m and by the specific form of the function $P_i(k)$.

Let us place a number of projectiles n=10, a number of targets m=5.

Functions $P_i(k)$ let us assign as follows:

$$\left.\begin{aligned}
&P_1(k)=1-0.2^k,\\
&P_2(k)=1-2\cdot 0.6^k-0.2^k;\\
&P_3(k)=1-0.1^k,\\
&P_4(k)=1-0.7^k-0.5^k+0.2^k;\\
&P_5(k)=1-0.5^k.
\end{aligned}\right| \qquad (14.13)$$

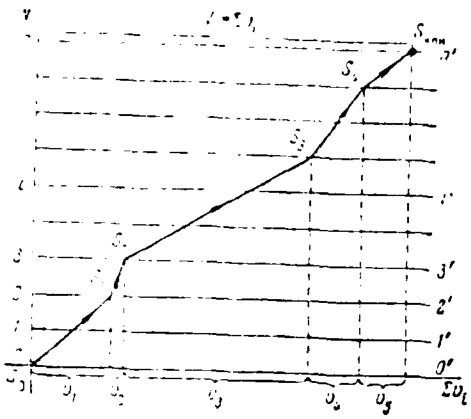


Fig. 14.1.

Page 141.

Let us note that from preset five functions two P_2 and P_4) are converted into zero not only with k=0, but also with k=1, i.e., the second and fourth purposes possess considerably smaller, in comparison with the first, the third and the fifth, by vulnerability; to these targets for us it is necessary to select no less than on the basis of two projectiles:

$$k_2 \geqslant 2, \qquad k_4 \geqslant 2. \qquad (14.14)$$

Let us pass from functions (14.13) to the new functions

$$v_i(k) = |\ln P_i(k)| \qquad (i = 1, 2, \ldots, 5) \qquad (14.15)$$

and for simplification in further calculations let us make table the values of all functions $v_i(k)$ (see Table 14.1).

Plotted function $v_i(k)$ $(i=1, \ldots, 5)$ they are represented in Fig. 14.2. Since functions are determined only milking the integer values of argument k, then lines Fig. 14.2 shows curves, but by broken lines.

Process of planning/gliding let us develop on the standard diagram. As the value, which characterizes the issue of the i step/pitch, we will examine number n_i of the projectiles, which remained at our disposal after the extraction of means to the i-th target.

1. We fix/record issue of fourth step/pitch n_4 - number of projectiles, which remained at our disposal after extraction of projectiles to first four purposes. Let us determine the borders in which it can be located by n_4. To fifth target it is necessary to leave not less than one projectile:

$$n_4 \geqslant 1.$$

Table 14.1.

k	$v_1(k)$	$v_2(k)$	$v_3(k)$	$v_4(k)$	$v_5(k)$
0	∞	∞	∞	∞	∞
1	0.223	∞	0,105	∞	0,693
2	0,041	1.139	0,010	1,204	0,288
3	0,008	0,552	0,001	0,616	0,134
4	0,002	0,298	0,000	0,358	0,065
5	0,000	0.169	0,000	0,222	0,032
6	0,000	0,098	0,000	0,143	0,016
7	0,000	0,058	0,000	0,094	0,008
8	0,000	0,034	0,000	0,063	0,004
9	0,000	0,020	0,000	0,043	0,002
10	0,000	0,012	0,000	0,030	0,001

Page 142.

On the first four targets were released not less 1+2+1+2=6 projectiles, remained not more than four. Thus,

$$1 \leqslant n_4 \leqslant 4.$$

It is obvious, optimum control on the fifth step/pitch lies in the fact that all remaining n_4 projectiles to isolate into the fifth purpose:

$$k_5^*(n_4) = n_4. \qquad (14.16)$$

In this case the value of criterion V, reached at the fifth step/pitch, is converted into the minimum and it is equal

$$V_5^*(n_4) = v_5(n_4). \qquad (14.17)$$

This value for any value of n_4 can be found from the latter/last column of Table 14.1. Let us register in new table 14.2 conditional optimum control on the fifth step/pitch $k_5^*(n_4)$ and value of criterion $V_5^*(n_4)$, reached at this control at the fifth step/pitch.

Table 14.2 is contained by the task of optimization at the fifth step/pitch. Let us note that here it does not have sense to construct graphs, since the possible values of argument n_4 integral and there are few of them.

2. Let us set n_3 - number of projectiles, which remained after extraction of resources to first three purposes.

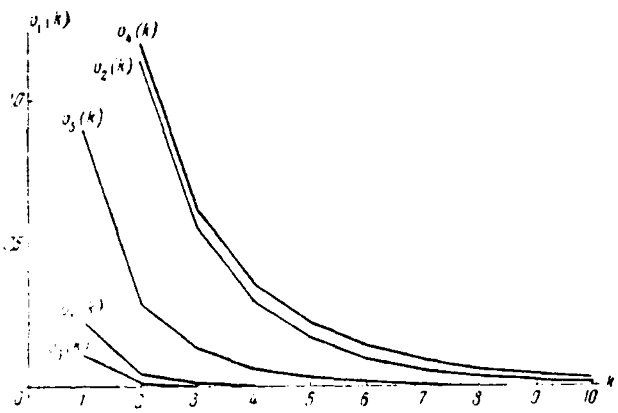


Fig. 14.2.

Page 143.

Let us determine the boundaries in which lies/rests n_3. To the remaining two targets - fourth and the fifth - it is necessary to reserve not less than three projectiles, hence

$$n_3 \geqslant 3;$$

on the other hand, to the first three purposes is spent not less 1+2+1=4 projectiles; remained not more than six. Thus,

$$3 \leqslant n_3 \leqslant 6.$$

According to standard procedure to us is necessary with each of the possible values of n_3 to minimize "prize" $V^+_{4,5}$ for the latter/last two steps/pitches during any control at the next-to-last step/pitch and optimum control on the latter:

$$V^+_{4,5}(n_3, k_4) = v_4(k_4) + V^*_5(n_4).$$

or, taking into account that $n_4=n_3-k_4$,

$$V_{4,5}^{+}(n_3,\ k_4)=v_4(k_4)+V_5^{*}(n_3-k_4). \qquad (14.18)$$

The conditional optimum prize $V_{4,5}^{*}(n_3)$ and conditional optimum equation $k_4^{*}(n_3)$ they will be located from the condition

$$V_{4,5}^{*}(n_3)=\min_{k_4}\{V_{4,5}^{+}(n_3,\ k_4)\}. \qquad (14.19)$$

where the minimum is taken in terms of all values of a number of chosen projectiles k_4, permitted with this n_3.

Using data of Tables 14.1 and 14.2 (into the latter it is necessary to enter with n_3-k_4 instead of n_4) let us make table from two inputs for function (14.18) (see Table 14.3).

The drawn a line graphs/counts Table 14.3 correspond to the impossible, with this n_3, values k_4. In each table row is emphasized minimum value.

Table 14.2.

n_4	$k_4^*(n_4)$	$V_4^*(n_4)$
1	1	0,693
2	2	0,288
3	3	0,134
4	4	0,065

Page 144.

It equally $V^*_{4,5}(n_3)$ determines conditional optimum control at the fourth step/pitch $k_4*(n_3)$. Let us reduce these data in Table 14.3.

Table 14.4 solves completely the problem of the optimization of the fourth step/pitch.

3. We optimize third step/pitch. We are assigned by values of n_2 for which we define the boundaries. For the first two steps/pitches is spent not less than three projectiles; remained not more than seven.

To the latter/last three steps/pitches it is must be reserved not less than four projectiles; consequently,

$$4 \leqslant n_2 \leqslant 7.$$

We construct table with two input of the values of the function

$$V^*_{3,4,5}(n_2, k_3) = v_3(k_3) + V^*_{4,5}(n_2 - k_3) \quad (14.20)$$

(Table 14.5) and we seek in each line of this table minimum number:

$$V^{*}_{3,4,5}(n_2)=\min_{k_3}\{V^{*}_{3,4,5}(n_2, k_3)\}. \qquad (14.21)$$

Table 14.3.

n_3 \ k_4	2	3	4	5
3	1,897	—	—	—
4	[illegible]	1,309	—	—
5	[illegible]	0,904	1,061	—
6	[illegible]	0,750	0,646	0,915

Table 14.4.

n_3	$k_4^*(n_3)$	$V_{4,5}^*(n_3)$
3	2	1,897
4	3	1,309
5	3	0,904
6	4	0,646

Page 145.

We stress in each to line a minimum number, find $k_3^*(n_2)$ and $V_{3,4,5}^*(n_2)$ and we write/record them in Table 14.6.

The optimization of the third step/pitch by these is exhausted.

We analogously optimize the second step/pitch. we find the borders of values n_1: $6 \leqslant n_1 \leqslant 9.$

We construct table with two input of the values of the function

$$V_{2,3,4,5}^{+}(n_1, k_2) = v_2(k_2) + V_{3,4,5}^*(n_1 - k_2) \quad (14.22)$$

(see Table 14.7).

In each line of this table we find the minimum number:

$$V^{*}_{2,3,4,5}(n_1) = \min_{k_2} \{V^{*}_{2,3,4,5}(n_1, k_2)\}. \qquad (14.23)$$

We further construct table for the optimization of the second step/pitch (Table 14.8).

Table 14.5.

k_3 / n_2	1	2	3	4
4	2,002	—	—	—
5	1,414	1,907	—	—
6	1,009	1,319	1,898	—
7	0,751	0,914	1,310	1,897

Table 14.6.

n_2	$k_3^*(n_2)$	$V^*_{3,4,5}(n_2)$
4	1	2,002
5	1	1,414
6	1	1,009
7	1	0,751

Page 146.

4. We optimize first step/pitch; value $n_0=10$ preset and is not varied; therefore simply we seek minimum of k_1 of function

$$V^+_{1,2,3,4,5}(10,\ k_1)=v_1(k_1)+V^*_{2,3,4,5}(10-k_1). \quad (14.24)$$

being absolute minimum of criterion V:

$$V^*=V^*_{1,2,3,4,5}=\min_{k_1}\{V^+_{1,2,3,4,5}(10,\ k_1)\}. \quad (14.25)$$

The values of function (14.24) are given in table 14.9.

Smallest of numbers in table 14.9 is equal

$$V^*=V^*_{1,2,3,4,5}=1.784$$

and it is achieved by optimum control at the first step/pitch

$$k_1^*=1.$$

Table 14.7.

n_1 \ k_2	2	3	4	5
6	3.141	—	—	—
7	2.553	2.554	—	—
8	2.148	1.966	2.300	—
9	1.890	1.561	1.712	2.171

Table 14.8.

n_1	$k_2^*(n_1)$	$V^*_{2,3,4,5}(n_1)$
6	2	3.141
7	2	2.553
8	3	1.966
9	3	1.561

Table 14.9

k_1	1	2	3	4
$V^*(10, k_1)$	1.784	2.007	2.561	3.143

Page 147.

In this case probability W of the destruction of all targets reaches its maximum:

$$W^* = e^{-V^*} \approx 0.168.$$

5. Passing again entire sequence of stages from beginning toward the end, we find unconditional optimum control on each step/pitch, beginning from the first:

$$k_1^* = 1; \quad n_1^* = 10 - k_1^* = 9.$$

included in Table 14.8 with $n_1=9$, we find

$$k_2^* = k_2^*(9) = 3; \quad V^*_{2,3,4,5} = 1.561.$$

Further we have

$$n_2^* = 9 - 3 = 6.$$

included in Table 14.6 with u_2=6, let us find

$$k_3^* = k_3^*(6) = 1; \quad V_{3,4,5}^* = 1.009.$$

Further

$$n_3^* = 6 - 1 = 5;$$

from Table 14.4 it has

$$k_4^* = k_4^*(5) = 3; \quad V_{4,5}^* = 0.904.$$

Further,

$$n_4^* = 5 - 3 = 2;$$

from Table 14.2

$$k_5^* = k_5^*(2) = 2; \quad V_5^* = 0.288.$$

Thus, optimum control is found:

$$U^* = (1, 3, 1, 3, 2),$$

i.e., for achievement of the maximum kill probability of all targets W it is necessary to isolate into the first and by third of target on one projectile, to the second and fourth purposes - on three projectiles, and the remaining two projectiles to isolate into the fifth purpose.

For the construction of trajectory in the phase space it is necessary to still compute values v_i^* (i=1, 2, 3, 4, 5) the "gains" ["prizes"], reached in the single stages at the optimum control.

Page 148.

We have:

$$
\begin{aligned}
v_1^* &= V_{1,2,3,4,5}^* - V_{2,3,4,5}^* = 1.784 - 1.561 = 0.223;\\
v_2^* &= V_{2,3,4,5}^* - V_{3,4,5}^* = 1.561 - 1.009 = 0.552;\\
v_3^* &= V_{3,4,5}^* - V_{4,5}^* = 1.009 - 0.904 = 0.105;\\
v_4^* &= V_{4,5}^* - V_5^* = 0.904 - 0.288 = 0.616;\\
v_5^* &= V_5^* = 0.288.
\end{aligned}
$$

The optimum trajectory of point S in the phase space will appear, as shown in Fig. 14.3.

Let us note that in this elementary example is demonstrated not the most economical method of the solution of the problem: perhaps, simpler it would be solve the task of rational control not of the dynamic programming, but simple the countershaft (number of possible versions it is not too great). However, in the more complex problems the advantages of the method of dynamic programming become evident.

b) the distribution of resources for increasing the reliability of technical device/equipment. Let there be the technical device/equipment A, which consists of m aggregates/units, or the assemblies

$$A_1, A_2, \ldots, A_m$$

(Fig. 14.4).

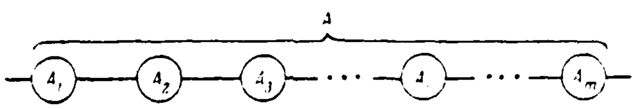


Fig. 14.4.

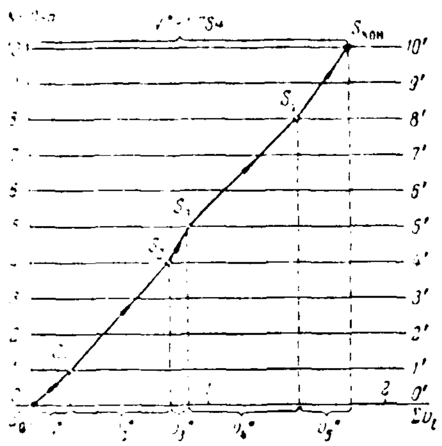


Fig. 14.3.

The failure-free operation of each of the assemblies is necessary for the work of device/equipment A as a whole. Aggregates/units can go out of order, moreover independently of each other. the reliability (probability of failure-free operation) of entire device/equipment is equal to the product of the reliability of the single assemblies

$$W = \prod_{i=1}^{n} p_i. \qquad (14.26)$$

where p_i — reliability of the i assembly.

For increasing the reliability of entire device/equipment is isolated some sum of resources a_0. These rescurces (expressed in the money, the weights or otherwise) can be in an arbitrary manner distributed between the actions, which raise the reliability of single assemblies. In order to increase the reliability of the i assembly from p_i to $P > p_i$ it is necessary to expend the sum, equal to $f_i(P, p_i)$.

It is necessary so to distribute the tempered resources in order to do reliability of entire device/equipment of maximum.

This task, and previous, also has multiplicative criterion, on it differs from it in terms of the fact that the control carries not discrete/digital, but continuous character and consists of the extraction to each assembly (stage) of the specific sum of resources x_i (i=1, 2, ..., m). After the conversion of criterion W to the additive form the logarithmic operation before us will be the ordinary task of distributing the resources/lifetimes with the redundancy (moreover in its simplest form when in each stage are expended/consumed to the end/lead all allocated resources), with the difference that the "income" is not maximized, but it is minimized.

Page 150.

We recommend to reader as the control to solve the task of distributing the resources to an increase in the reliability with following concrete/specific/actual data

$$m = 5;\ p_1 = 0.90;\ p_2 = 0.95;\ p_3 = 0.93;$$
$$p_4 = 0.85;\ p_5 = 0.97.$$
$$Z_0 = 4.$$
$$f_1(P,\ p_1) = 2\left|\ln\frac{1-P}{1-p_1}\right|;$$
$$f_2(P,\ p_2) = 0.8\left|\ln\frac{1-P}{1-p_2}\right|;$$
$$f_3(P,\ p_3) = 0.5\left|\ln\frac{1-P}{1-p_3}\right|;$$
$$f_4(P,\ p_4) = 0.3\left|\ln\frac{1-P}{1-p_4}\right|;$$
$$f_5(P,\ p_5) = 3\left|\ln\frac{1-P}{1-p_5}\right|.$$

All functions $f_i(P)$ are determined only for $P > p_i$ (i=1, ..., 5).

§15. Stochastic tasks of dynamic programming.

In practice frequently meet such tasks of the planning/glidings in which noticeable role play the random factors, which affect both the state of system S and prize W. In such problems, the controlled process is not completely determined by the initial state S_0 and the selected control U, while to a certain degree it depends on the case. Let us agree also the task of planning/gliding to call "stochastic" (probabilistic).

Certain representation about similar tasks we already obtained in §12, where was found out the optimum distribution of weapons of destruction according to the defended targets. Depending on the case a number of surviving to this nuclear weapons of destruction could be the fact, etc., i.e., the state of system, in the essence, was random.

Page 151.

However, with the solution of this problem we were founded to the examination only of the average/mean characteristics of process, i.e., they solved its not as stochastic, but as ordinary problem of dynamic programming.

This method, which is reduced to the fact that the random process previously, even before the solution of task, is replaced by its averaged, not random, determined model, is approximate and will use far not always. It gives fair results only in those tasks of which the controlled system consists of the sufficiently multiple objects (as in §12: aircraft, instruments), and despite the fact that the state of each of them varies randomly, in the mass these randomness mutually are liquidated, they are levelled.

In many stochastic tasks of planning/gliding this method cannot be used; into some it gives too great errors, in others and it is completely impossible. And, in any case, does always arise the question: strongly whether will be changed optimum control, if we disregard/neglect randomness and to replace stochastic task of that determined? In order to answer this question, it is necessary to be able to solve stochastic problems of dynamic programming taking into account random factors.

In the present paragraph we will give fundamental approach to such tasks and let us plan general/common/total outline of the solution. In the following §16 we will in detail dismantle/select a specific example.

The general/common/total formulation stochastic problem of dynamic programming can be described as follows.

Let there be the physical system S, which in the course of time varies its state. We can to a certain degree act on this process, directing him to the desired side, but we monitor this process not completely, since its course, besides the control, depends also on the case. This process we will call the "random controlled process".

Let us assume that with the course of the process is connected our some interest, which is expressed by criterion ("prize") W, which to us it is desirable to become maximum.

Page 152.

Criterion W is additive:

$$W = \sum_{i=1}^{m} w_i. \qquad (15.1)$$

where w_i — the "gain", acquired in the i stage of process.

Since the state of system S is by chance, then random proves to be gain w_i in each stage, and total gain W.

We would like to select such control U, during which prize W would be converted into the maximum. But can we this do? it is obvious, no: during our any control prize W will remain random. However, we can select such control during which the average/mean value of the random prize W will be maximum. Let us designate the average/mean value (mathematical expectation) of value W by the letter $\overline{W}$:

$$\overline{W} = M[W]. \qquad (15.2)$$

Taking into account formula (15. 1) and using the property of mathematical expectation[1], let us register $\overline{W}$ in the form

$$\overline{W} = \sum_{i=1}^{m} \overline{w}_i. \qquad (15.3)$$

where $\overline{w}_i$ — average/mean prize in the i stage.

FOOTNOTE [1]. The mathematical expectation of sum of random variables is equal to the sum of their mathematical expectations. ENDFOOTNOTE.

Thus, in stochastic tasks instead of the criterion itself W which is accidental, is examined its average/mean value $\overline{W}$; this criterion is also additive.

The task of dynamic programming is reduced to the following: to select this optimum control u^*, which consists of optimum controls $U_1^*, U_2^*, \ldots, U_m^*$ on the single stages so that the additive criterion $\overline{W}$ would become maximum.

It would seem, matter is reduced to the simple replacement of criterion; however this not ends. The difference between stochastic and determined diagrams of dynamic programming is much deeper: it concerns the very structure of optimum control.

Actually/really, let us recall the overall diagram of dynamic programming in the determined processes, without the participation of randomness (§7). It consists of the following.

Page 153.

Is fixed/recorded the state of system S_{i-1} afterward (i-1)-th step/pitch, and for each of such states is sought conditional optimum control at the i step/pitch. In this case the state of system after i step/pitch S_i is completely determined by previous state S_{i-1} and used at the i step/pitch control U_i:

$$S_i = S_i(S_{i-1}, U_i). \qquad (15.4)$$

Equally prize w_i at the i step/pitch is completely determined by the state of system S_{i-1} in the beginning of this step/pitch and by used control U_i:

$$w_i = w_i(S_{i-1}, U_i). \qquad (15.5)$$

So whether this will be in stochastic tasks? No, not thus. The state of system S_i after the i step/pitch is not completely determined by state S_{i-1} and control U_i, but it depends also on the case. State S_i with given ones S_{i-1} and U_i is random, and on given ones S_{i-1} and U_i depends only probability distribution for different versions of state S_i.

Gain w_i at the i step/pitch is not also completely determined by the previous state of system S_{i-1} and by used control U_i, but it is random variable, and on S_{i-1} and U_i depends only probability distribution between its possible values. But that as us interests not random gain itself w_i, but only its average/mean value at each step/pitch, this random variable it is possible to average taking into account probability distribution and to introduce into the examination conditional mean prize at the i step/pitch in preset state S_{i-1} after (i-1)-th step/pitch and the specific control at i step/pitch U_i:

$$\bar{w}_i(S_{i-1}, U_i). \qquad (15.6)$$

We will assume that both the probability distribution for random state S_i and conditional mean gain (15.6) they depend only on S_{i-1} and U_i do not depend on the "prehistory" of process, i.e., from that how, when and as a result of what control system arrived into the state S_{i-1} [1].

FOOTNOTE [1]. In other words, the controlled process has Markov character. ENDFOOTNOTE.

Page 154.

By our task will determine for each of the possible random issues any step/pitch conditional optimum control at the following step/pitch. Thus, in stochastic diagram optimum control itself U* will be random and will be each time realized differently, depending on that how will be developed random process. The matter of that planning/gliding - on to work out the fixed program of control, and to indicate for each step/pitch the control which one should answer any random issue of the previous step/pitch.

In this - the basic difference between the determined and stochastic tasks of dynamic programming. In the determined task optimum control is single and is indicated previously as the wired program of action. In stochastic task optimum control is random and is chosen in the course of process itself depending on the randomly established situation. This is - control with the "feedback" from the actual state of system to the control.

Let us pay attention to one more circumstance. In the determined diagram, passing process in stages from the end/lead at the beginning, we on each stage also found a whole series of conditional optimum controls, but of these all controls in the final analysis was realized only one. In stochastic diagram is not thus. Each of the conditional optimum controls can prove to be actually realizable, if the previous course of random process puts system into the appropriate state.

But how to find conditional optimum control [illegible] on the i step/pitch of stochastic process? This is - such control which, being used at the i step/pitch, converts into the maximum conditional mean prize at all subsequent steps/pitches: from the i-th to the m-th inclusively. How to find this maximum? In perfect analogy how we made in the determined diagram, with the difference that instead of the prize itself which was accidental, is examined its average/mean value.

Page 155.

The process of planning/guiding, as always, is turned/run up, beginning from the latter/last (m-th) step/pitch. Is fixed/recorded the state of system S_{m-1} after (m-1)-th step/pitch, and under this condition is sought the control $U_m^*(S_{m-1})$, which converts into the maximum average/mean conditional prize $\overline{w}_m(S_{m-1})$ at the m step/pitch:

$$\overline{W}_m^*(S_{m-1})=\max_{U_m}\{\overline{w}_m(S_{m-1},\ U_m)\}. \qquad (15.7)$$

When is found dependence $U_m^*(S_{m-1})$ and $\overline{W}_m^*(S_{m-1})$ on S_{m-1}, the optimization of latter/last step/pitch is completed. In whatever state S_{m-1} randomly arrived the system after m-1 steps/pitches, we already know that to us to make further and what maximum average/mean prize we will obtain at the latter/last step/pitch.

Then is optimized (m-1)-th step/pitch. Here matter proceeds so not simply. Let us first fix state S_{m-2} after m-2 steps/pitches and will find average/mean prize $\overline{W}_{m-1,m}^*(S_{m-2},\ U_{m-1})$ on two latter/last steps/pitches during any control U_{m-1} at (m-1)-th step/pitch and during the optimum control on the latter which to us it is already known. How to find this average/mean prize? We will discuss as follows. During any preset S_{m-2} and control U_{m-1} state S_{m-1} will be random, but we know for it the probability distribution, which depends on S_{m-2} and U_{m-1}. Random state S_{m-1} determines by itself maximum average/mean conditional prize at m step/pitch $\overline{W}_m^*(S_{m-1})$:

average this value taking into account the probability distribution of state S_{m-1}. We will obtain the "twice averaged" maximum conditional prize at the m step/patch, which depends no longer on S_{m-1} (on it is done the second averaging), but on S_{m-2} and U_{m-1}. Let us designate this prize

$$\overline{\overline{W}}^{*}_{m}(S_{m-2}, U_{m-1}). \qquad (15.8)$$

To it it is necessary to adjoin average/mean conditional gain at (m-1)-th step/pitch during control U_{m-1} at this step/pitch:

$$\overline{w}_{m-1}(S_{m-2}, U_{m-1}); \qquad (15.9)$$

we will obtain

$$W^{*}_{m-1,m}(S_{m-2}, U_{m-1}) = \\ = \overline{w}_{m-1}(S_{m-2}, U_{m-1}) + \overline{\overline{W}}^{*}_{m}(S_{m-2}, U_{m-1}). \qquad (15.10)$$

Page 156.

Let us find the control on (m-1)-th step/pitch with which value (15.10) is converted into the maximum:

$$W^{*}_{m-1,m}(S_{m-2}) = \max_{U_{m-1}} \{W^{*}_{m-1,m}(S_{m-2}, U_{m-1})\} = \\ = \max_{U_{m-1}} \{\overline{w}_{m-1}(S_{m-2}, U_{m-1}) + \overline{\overline{W}}^{*}(S_{m-2}, U_{m-1})\}. \qquad (15.11)$$

Let us designate this conditional optimum control $U^{*}_{m-1}(S_{m-2})$. Thus, as a result of optimization of (m-1)-th step/pitch are found conditional optimum control $U^{*}_{m-1}(S_{m-2})$ and corresponding to it maximum average/mean conditional prize on two latter/last steps/pitches $W^{*}_{m-1,m}(S_{m-2})$.

Analogously is optimized any i step/pitch. For each issue of i step/pitch S_i is already known conditional maximum average/mean prize at the subsequent steps/pitches:

$$\bar{W}^*_{i+1, \ldots, m}(S_i). \quad (15.12)$$

Since state S_i is random and its distribution depends on S_{i-1} and U_i, then it is possible random gain (15.12) to average taking into account probability distribution S_i and to obtain the "twice averaged" gain

$$\bar{\bar{W}}^*_{i+1, \ldots, m}(S_{i-1}, U_i).$$

After forming it with the average/mean prize at the i step/pitch during any control U_i we will obtain

$$\bar{W}^*_{i, \ldots, m}(S_{i-1}, U_i) = = \bar{w}_i(S_{i-1}, U_i) + \bar{\bar{W}}^*_{i+1, \ldots, m}(S_{i-1}, U_i). \quad (15.13)$$

Random optimum control on i step/pitch $U_i^*(S_{i-1})$ will be located as the control during which value (15.13) reaches the maximum:

$$W^*_{i, \ldots, m}(S_{i-1}) = \max_{U_i} \{\bar{W}^*_{i, i+1, \ldots, m}(S_{i-1}, U_i)\} = = \max_{U_i} \{\bar{w}_i(S_{i-1}, U_i) + \bar{\bar{W}}^*_{i+1, \ldots, m}(S_{i-1}, U_i)\}. \quad (15.14)$$

Applying formula (15.14) consecutively/serially at each step/pitch, let us find conditional optimum control and maximum conditional mean prize on all steps/pitches of process to the second inclusively.

Page 157.

The optimization of the first step/pitch has some special features/peculiarities, connected with the fact that the initial state S_0, as a rule, random is not and it must not be varied taking into account probability distribution[1].

FOOTNOTE [1]. For simplicity of presentation we take only that case when S_0 in the accuracy known. ENDFOOTNOTE.

Since S_0 is not by chance, then not random (completely determined) will be the optimum control u^*_1 during the first stage, and average/mean prize taking into account this stage will be

$$\overline{W}^* = \overline{\overline{W}}^*_{1, 2, \dots, m}. \quad (13.15)$$

Thus, the process of dynamic programming is completed; is found the optimum control

$$U^* = (U^*_1, U^*_2(S_1), U^*_3(S_2), \dots, U^*_m(S_{m-1})), \quad (13.16)$$

all elements/cells of which, except the first, are random and depend on the state of system. Corresponding to this control maximum average/mean prize is equal to (13.15).

Recommendations regarding the use/application of this control are given in the following form: at the first step/pitch to apply control U^*_1 and to await the results of this step/pitch; on the second, depending on issue S_1 of the first step/pitch, to choose the optimum control $U^*_2(S_1)$ and so forth.

Page 158.

§16. Example stochastic task of the dynamic programming: the combined fire control and of exploration.

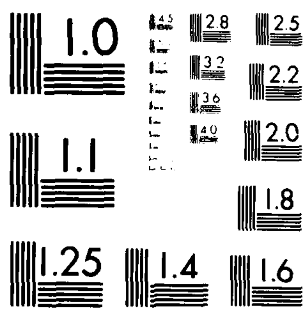
Let us consider an elementary example stochastic task of dynamic programming.

Is produced shooting at some target Ts. In our disposal there are by m of projectiles; for the destruction of target it is sufficient one incidence/impingement. Shots are not depended; the hit probability into the target with each shot is equal to p. Each projectile has considerable cost/value, and therefore to undesirably expend projectiles for nothing, on the already affected target. We have the capability, if we wish, after each shot to produce exploration (for example, to send reconnaissance aircraft) and to establish/install, is affected target or not: if it is affected, we will cease shooting and part of the projectiles will be saved. However, to too frequently send exploration is not advantageous, since this dearly is bypassed, but direct damage to target in this

AD-A098 406 FOREIGN TECHNOLOGY DIV WRIGHT-PATTERSON AFB OH F/G 12/1
ELEMENTS OF DYNAMIC PROGRAMMING.(U)
FEB 81 Y S VENTTSEL'
UNCLASSIFIED FTD-ID(RS)T-1515-80 NL

4 of 4
AD
A098406

END
DATE
FILMED
5 81
DTIC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

case we will not plot. Arises the question about the reasonable control of the process of the bombardment of target and reconnaissance operations.

It is obvious, the solution must depend on cost of projectile, the cost/value of exploration, and also from the value of target. Let the cost/value of projectile s; the cost/value of one exploration r; in the case of the destruction of target we obtain the "premium" of A (expressing, for example, by the cost/value of target itself, either by the cost/value of the reducing work of the opponent, or finally by our material damage which it was possible to prevent, after striking target). Our task - of planning bombardment and exploration so as to become maximum pure/clean "income" from the entire operation ("incomes" from the expenditures of projectiles and send operation of prospectors it goes without saying are considered negative). To select optimum control - this means to indicate, when (after what in order of firings) it is necessary to send exploration and when to cease shooting.

The physical system, which we manage, consists of the informational component/link (prospector) and weapons of destruction (projectiles). The process, which takes place in the system, is the random controlled process, and the task of optimization must be solved on stochastic diagram. Criterion is average/mean "income" $\bar{W}$

from the entire series of measures (shooting, exploration).

In this example we deal concerning the discrete/digital random controlled controlled process in course of which system S irregularlyly passes of one state to another. Let us determine the possible states of system as follows.

Page 159.

We will indicate that system S is in state $S^{(k)}$ $(0 < k \leqslant m-1)$, if shooting at the target is not yet completed, but from the moment/torque of obtaining the latter/last information about the target is produced exactly k of shots [1].

FOOTNOTE [1]. Let us note that the information about the target can consist only of the fact that the "target is not affected", since otherwise shooting would be immediately ended. ENDFOOTNOTE.

For example, "state $S^{(0)}$" it indicates: is recently produced the exploration, which discovered that the target is not affected, and shooting is continued. It is obvious, if shooting yet was not begun, system also is in state $S^{(0)}$, since authentically it is known that the target is not affected. "State $S^{(1)}$" means that from the moment/torque of the admission of the latter/last information about the target, consisting in the fact that the target is not affected, are produced three shots and is not yet made decision about the curtailment of shooting.

If shooting is ended at any reason (are spent all projectiles, either it is discovered, that the target is already affected, or is simply accepted the solution to cease shooting), then we will indicate that the system is in state $S^{(m)}$.

By letters $S^{(k)}$ by superscripts we designate different possible states of system. One ought not to confuse them with designation S_i, which designates the state of system after i-th step/pitch. In our process the number of step/pitch will not at all coincide with the number of state.

As the phase space let us consider points $S^{(0)}, S^{(1)}, \ldots, S^{(m-1)}, S^{(m)}$ on the axis of abscissas (Fig. 16.1).

The transition of system abruptlyly from one state into another we will represent as arrow/pointer, is shown in Fig. 16.1.

Let us divide process into the steps/pitches. Natural "step/pitch" in this case is "shot". The fact that in actuality can be carried out not all m of shots, must not confuse us: always it is possible the afterward latter/last actually done shot to count off mentally still several fictitious ones, so that the total number of shots (real and fictiticus) would be always equal to m.

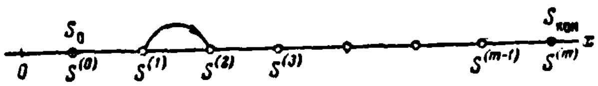


Fig. 16.1.

Page 160.

"Exploration" we will not count for the single step/pitch, since the otherwise total number of steps/pitches would not be fixed/recorded. If after this shot is produced exploration, then we will carry it to the same step/pitch, as shot.

Thus, we have three varieties of the steps/pitches: shot, shot with the subsequent exploration and fictitious shot.

Let us consider possible controls at each step/pitch. If system is in state $S^{(m)}$ - shooting is already ended, then there is no selection, and the only possible (it and optimum) control will be: not to shoot. But if system is in state $S^{(k)}$ $(0 \leqslant k \leqslant m-1)$, then at our disposal there are three versions of the control:

$U^{(1)}$ — to do first-order step/pitch, i.e., to shoot and not to send exploration,

$U^{(2)}$ — to do second-order step/pitch, i.e., to shoot and to send exploration,

$U^{(3)}$ — to do third-order step/pitch, i.e., cease shooting and to produce fictiticus shot.

Let us agree to consider that if the exploration, execution at the previous step/pitch, brought information "target it was affected", then shooting immediately ceases, i.e., control $U^{(3)}$ at the following step/pitch is forced.

Let us describe the motion of point S, which represents the state of system, in the phase space (Fig. 16.1). The initial state of system S_0 is completely determined coincides with point $S^{(0)}$ Final state $S_{кон}$ of point is completely determined coincides with point $S^{(m)}$. However, the intermediate positions of point S can be different, depending on the used control and success of shooting. For example, Fig. 16.2 depicts this concrete/specific/actual realization of the random process: first are done three shots without the exploration; point S irregularlyly is moved from $S^{(0)}$ in $S^{(1)}$, $S^{(2)}$, $S^{(3)}$. After the third shot is produced the exploration; it is explained that the target is not affected, and point S is returned to state $S^{(0)}$.

Page 161.

Then is produced four additional shots (point it is moved in $S^{(1)}, S^{(2)}, S^{(3)}, S^{(4)}$). it is sent exploration, it discovers that the target is affected, and shooting ceases (point S it jumps from $S^{(4)}$ in $S^{(m)}$). after which it is continued the reading of fictitious shots, if they yet not all are spent.

Each trajectory, similar to that to recently depicted, consists of n jumps (displacements/movements) of the phase space, moreover to all fictitious shots, produced already in state $S^{(m)}$. correspond "jumps on the spot" with the return to the same point $S^{(m)}$. After each step/pitch, at which was used control $U^{(2)}$ (shot with the subsequent exploration), point S can only either return in $S^{(n)}$ (if exploration it communicated: "target was not affected"), or pass in $S^{(m)}$ (if exploration it communicated: "target was affected", and is made the decision to cease shooting). Let us note that the solution to cease shooting to completely thoughtlessly accept afterward Proceedings "target it is not affected", since, keeping in mind to cease shooting with any report of exploration, it would be not more reasonably completely send it and not produce the senseless expenditures r.

Before us will cost the task - of selecting optimum control at each step/pitch, i.e., to indicate, which of three controls

$U^{(1)}, U^{(2)}, U^{(3)}$ must be applied at this step/pitch depending on the issue of the previous step/pitch. It is obvious, optimum control must depend on the parameters of the problem: number of projectiles m, hit probability p, cost/value of projectile s, cost/value of exploration r and the "premium" of A. Let us note that with some relationships/ratios of the parameters the problem is solved trivially.

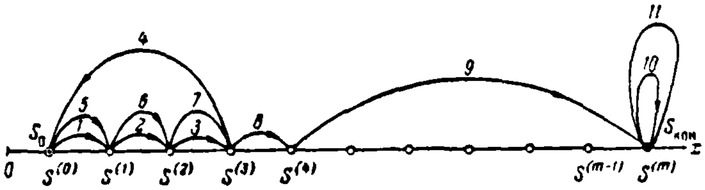


Fig. 16.2.

Page 162.

For example, if s>pA, i.e., the mean cost/value of one projectile is more than average/mean profit, which it yields with one shot, to shoot not at all has sense, i.e., at all steps/pitches it is necessary to apply control $U^{(3)}$. If s>A or r>ms, then it has sense to produce exploration and control $U^{(2)}$ must not be applied not at one step/pitch.

Let us consider the general case (at any values of the parameters) and will construct the diagram of the solution of problem by the method of dynamic programming.

First of all let us establish/install, from what states into what passes the system under the effect of specific controls $U^{(1)}$, $U^{(2)}$, $U^{(3)}$ and what in this case it is obtained average/mean prize.

From the state $S^{(m)}$ the system cannot pass anywhere else: single possible (it and optimum) control is $U^{(3)}$: prize w_i (and it goes without saying average/mean prize $\bar{w}_i$) in this case at any i-th step/pitch is equal to zero:

$$\bar{w}_i(S^{(m)}) = w_i(S^{(m)}) = 0. \qquad (16.1)$$

If system is in state $S^{(k)}$ $(0 \leqslant k \leqslant m-1)$, its further behavior and prize at i-th step/pitch depend on control.

Let us assume that is used control $U^{(1)}$ (shot without the exploration). System passes from $S^{(k)}$ to following in order state $S^{(k+1)}$, unless ended all projectiles; if on the next shot ended projectiles, system passes into state $S^{(m)}$.

The prize, which we in this case will obtain, does not depend on that, passed system in $S^{(k+1)}$ or in $S^{(m)}$. It consists of two components/terms/addends: not random and random. Not random component/term/addend is negative and is equal - s (we for sure lose with the shot the cost/value of projectile). Random component/term/addend is positively and the premium which we on this shot can obtain, but we can and not obtain. Premium will be obtained at this step/pitch, if the corresponding shot proves to be successful, will strike target. Let us find the probability of this event.

Page 163.

So that the target would be affected by the data by shot, it is necessary, in the first place, so that it would not be affected by previous k by the shots, released since would be obtained information "target it was not affected"; furthermore, it must be affected by the precisely this shot. The probability of this set of events is equal to

$$(1-p)^k p.$$

With this probability this shot will bring to us "income" in the form of premium A; with probability $1-(1-p)^k p$ it will not bring to us income, i.e., will bring "income". Average value A and 0 taking into account their probabilities; we will obtain average/mean income from this shot

$$A(1-p)^k p+0\cdot[1-(1-p)^k p]=A(1-p)^k p.$$

Subtracting from it the cost/value of projectile s, we will obtain the full/total/complete average/mean prize, acquired on one i-th step/pitch in the initial state of system $S^{(k)}$ and control $U^{(1)}$:

$$\overline{w}_i(S^{(k)},\ U^{(1)})=A(1-p)^k p-s. \qquad (16.2)$$

Let us assume now that to the system, which is found in state $S^{(k)}$, is used control $U^{(2)}$ (shot with the exploration). Further fate of system depends on what information it will communicate exploration. If this information will be "target affected", then system will pass into state $S^{(0)}$; if "target is affected", then system will pass into state $S^{(m)}$. Probability cf first of these events is equal $(1-p)^{k+1}$; of the second $1-(1-p)^{k+1}$.

Let us count average/mean prize during the use/application of control $U^{(2)}$. It is composed of the reliable expenditure/consumption (cost/value of projectile and exploration), equal - (s+r), and the average/mean income, equal to $A(1-p)^k p$; altogether

$$\overline{w}_i(S^{(k)},\ U^{(2)}) = A(1-p)^k p - (s+r). \qquad (16.3)$$

Let us assume that to the system, which is found in state $S^{(k)}$, is used control $U^{(3)}$ (shooting it is ended). Point S will move in $S^{(m)}$, and prize (both actual and average/mean) will be equal to zero:

$$\overline{w}_i(S^{(k)},\ U^{(3)}) = 0. \qquad (16.4)$$

Page 164.

Relying on these data, we will construct the process of the optimization of control. Let us begin with m-th step/pitch. Let afterward (m-1) step/pitch the system be in state S_{m-1} (it is not necessary to mix this state with point $S^{(m-1)}$!). There can be two cases: either we have already been found at point $S^{(m)}$ (shooting it is ended):

$$S_{m-1} = S^{(m)},$$

or shooting is not yet ended, and then we can be found in any from the remaining points:

$$S_{m-1} = S^{(k)} \qquad (0 \leq k \leq m-1).$$

In the first case the only possible (and, obviously, optimum) will be control $U^{(3)}$:

$$U_m^*(S^{(m)}) = U^{(3)}. \quad (16.5)$$

In the second case possibly only of two controls: $U^{(1)}$ and $U^{(3)}$ (since to send exploration after latter/last shot no longer has a sense).

Let us find average/mean prize on the latter/last step/pitch for each of these controls. If we will use control $U^{(1)}$ (shot without the exploration), then average/mean prize at m-th step/pitch will be (see formula (16.2))

$$\bar{w}_m(S^{(k)}, U^{(1)}) = A(1-p)^k p - s \quad (k = 0, 1, \ldots, m-1). \quad (16.6)$$

If we will use control $U^{(3)}$ (let us cease shooting), then average/mean prize at m-th step/pitch will be

$$\bar{w}_m(S^{(k)}, U^{(3)}) = 0. \quad (16.7)$$

Conditional optimum control at m-th step/pitch $U_m^*(S^{(k)})$ will be that with which this average/mean prize reaches the maximum:

$$\bar{W}_m^*(S^{(k)}) = \max\{A(1-p)^k p - s;\ 0\}. \quad (16.8)$$

Thus, if

$$A(1-p)^k p - s > 0,$$

then it is necessary to choose at m-th step/pitch control $U^{(1)}$; but if

$$A(1-p)^k p - s < 0,$$

- control $U^{(3)}$. The optimization of latter/last step/pitch is completed.

Let us switch over to optimization (m-1) of step/pitch. Let afterward (m-2) step/pitch the system be in state S_{m-2}. Here again there can be two cases: either shooting is ended, i.e.

$$S_{m-2}=S^{(m)}.$$

or it is not yet ended:

$$S_{m-2}=S^{(k)} \qquad (0<k<m-2).$$

If $S_{m-2}=S^{(m)}$, again the only possible control is $U^{(3)}$, and maximum average/mean prize at two latter/last steps/pitches is equal to zero:

$$U^{*}_{m-1}(S^{(m)})=U^{(3)}; \qquad \overline{W}^{*}_{m-1,m}(S^{(m)})=0. \qquad (16.9)$$

Let us find conditional optimum control and conditional maximum average/mean prize on two latter/last steps/pitches when $S_{m-2}=S^{(k)}$ $(k\neq m)$. For this let us consider average/mean conditional prize $\overline{W}_{m-1,m}$ at two latter/last steps/pitches during any control at (m-1) the step/pitch and optimum control at m-th step/pitch. Let us look, in that is converted this prize during controls $U^{(1)}, U^{(2)}, U^{(3)}$ at (m-1) the step/pitch, and let us select among these three values great.

We know that the system after m-2 steps/pitches is in state $S^{(k)}$. Let us use to it to (m-1) step control $U^{(1)}$. In this case is uniquely determined the state of the system afterward (m-1) of the step/pitch:

$$S_{m-1}=S^{(k+1)},$$

and average/mean prize at (m-1) the step/pitch will be determined by formula (16.2). Therefcre we can immadiately write the expression

$$\overline{W}^{*}_{m-1,m}(S^{(k)},\ U^{(1)}) = \\ = \overline{w}_{m-1}(S^{(k)},\ U^{(1)}) + \overline{\overline{W}}{}^{*}_{m}(S^{(k)},\ U^{(1)}) = \\ = A(1-p)^{k}p - s + \overline{W}^{*}_{m}(S^{(k+1)}). \quad (16.10)$$

Page 166.

Let us assume that is applied control $U^{(2)}$. Prize $W^{*}_{m-1,m}$ will be composed of prize $\overline{w}_{m-1}$ at (m-1) the step/pitch (see formula (16.3)) and twice averaged maximum prize at latter/last step/pitch $\overline{\overline{W}}{}^{*}_{m}$. It we will find as follows. In preset state $S^{(k)}$ afterward (m-2) step/pitch and during control $U^{(2)}$ the system with probability $(1-p)^{k+1}$ will pass into state $S^{(0)}$, and then maximum average/mean prize at the latter/last step/pitch will be equal to $\overline{W}^{*}_{m}(S^{(0)})$; with probability $1-(1-p)^{k+1}$ it will pass in $S^{(m)}$, and then prize at the latter/last step/pitch will be equal to zero. Averaging these two prizes taking into account their probabilities, we will obtain

$$\overline{\overline{W}}{}^{*}_{m}(S^{(k)},\ U^{(2)}) = (1-p)^{k+1}\overline{W}^{*}(S^{(0)})$$

and

$$W^{*}_{m-1,m}(S^{(k)},\ U^{(2)}) = \\ = A(1-p)^{k}p - (s+r) + (1-p)^{k+1}\overline{W}^{*}(S^{(0)}). \quad (16.11)$$

Finally, during control $U^{(3)}$ at (m-1) the step/pitch average/mean prize for the latter/last two steps/pitches will be equal to zero:

$$\overline{W}^{*}_{m-1,m}(S^{(k)},\ U^{(3)}) = 0. \quad (16.12)$$

Let us find maximum of three values (16.10), (16.11) and (16.12):

$$\overline{W}^{*}_{m-1,m}(S^{(k)}) = \max \begin{Bmatrix} \overline{W}^{+}_{m-1,m}(S^{(k)},\ U^{(1)}), \\ \overline{W}^{+}_{m-1,m}(S^{(k)},\ U^{(2)}), \\ \overline{W}^{+}_{m-1,m}(S^{(k)},\ U^{(3)}) \end{Bmatrix} =$$

$$= \max \begin{Bmatrix} A(1-p)^{k}p - s + \overline{W}^{*}_{m}(S^{(k+1)}), \\ A(1-p)^{k}p - (s+r) + (1-p)^{k}\overline{W}^{*}_{m}(S^{(0)}), \\ 0 \end{Bmatrix}. \quad (16.13)$$

Then from controls $U^{(1)}, U^{(2)}, U^{(3)}$ at which reaches this maximum, and there is optimum conditional control $U^{*}_{m-1}(S^{(k)})$ at (m-1) the step/pitch.

Page 167.

Is in perfect analogy optimized any (i-th) step/pitch (1<i<m). Piksiruem isxod (i-1) step/pitch S_{i-1}. If

$$S_{i-1} = S^{(m)},$$

the further control will be $U^{(3)}$ and prize at all subsequent steps/pitches zero. If

$$S_{i-1} = S^{(k)} \qquad (0 \leqslant k \leqslant i-1),$$

the further behavior of system and average/mean prize depend on control. If to i-th step/pitch is used control $U^{(1)}$, then

$$S_i = S^{(k+1)},$$

and

$$\overline{W}^{+}_{i,i+1,\ldots,m}(S^{(k)},\ U^{(1)}) = \\ = \overline{w}_i(S^{(k)},\ U^{(1)}) + \overline{W}^{*}_{i+1,\ldots,m}(S^{(k)},\ U^{(1)}) = \\ = A(1-p)^{k}p - s + \overline{W}^{*}_{i+1,\ldots,m}(S^{(k+1)}). \quad (16.14)$$

Let us assume that at i-th step/pitch toward the system, which is found in state $S^{(k)}$ $(k \neq m)$ is used control $U^{(2)}$. Average/mean prize $\overline{W}^{+}_{i,i+1,\ldots,m}$ will be composed of average/mean prize $\overline{w}_i$ on i-th the step/pitch (see formula (16.3)) and twice averaged maximum prize $\bar{\bar{W}}^{*}_{i+1,\ldots,m}$ at all latter/last steps/pitches, beginning (i+1). This twice averaged prize let us find with the help of the following reasonings. In preset state $S^{(k)}$ afterward (i-1) step/pitch and during control $U^{(2)}$ at i-th step/pitch the system with probability $(1-p)^{k+1}$ will pass in $S^{(0)}$ and then maximum average/mean prize at the latter/last steps/pitches, beginning (i+1), there will be $\overline{W}^{*}_{i+1,\ldots,m}(S^{(0)})$ (this prize we already is found, optimizing (i+1) step/pitch); with probability $1-(1-p)^{k+1}$ the system will pass in $S^{(m)}$ and then average/mean prize at all subsequent staps/pitches will be equal to zero. Averaging these two prizes taking into account their probabilities, we will obtain

$$\bar{\bar{W}}^{*}_{i+1,\ldots,m}(S^{(k)},\ U^{(2)})=(1-p)^{k+1}\overline{W}^{*}_{i+1,\ldots,m}(S^{(0)})$$

and

$$\overline{W}^{+}_{i,i+1,\ldots,m}(S^{(k)},\ U^{(2)})=A(1-p)^{k}p-(s+r)+ +(1-p)^{k+1}\overline{W}^{*}_{i+1,\ldots,m}(S^{(0)}). \quad (16.15)$$

Page 168.

Finally, if to i-th step/pitch will be used control $U^{(3)}$, average/mean prize at all subsequent steps/pitches will be equal to zero:

$$W^{+}_{i,i+1,\ldots,m}(S^{(k)},\ U^{(3)})=0. \quad (16.16)$$

Let us find maximum of expressions (16.14), (16.15), (16.16):

$$\overline{W}^*_{i,i+1,\ldots,m}(S^{(k)}) = \max\begin{Bmatrix} \overline{W}^*_{i,i+1,\ldots,m}(S^{(k)},\ U^{(1)}), \\ \overline{W}^*_{i,i+1,\ldots,m}(S^{(k)},\ U^{(2)}), \\ \overline{W}^*_{i,i+1,\ldots,m}(S^{(k)},\ U^{(3)}) \end{Bmatrix} =$$

$$= \max\begin{Bmatrix} A(1-p)^k p - s + \overline{W}^*_{i+1,\ldots,m}(S^{(k+1)}), \\ A(1-p)^k p - (s+r) + (1-p)^{k+1}\overline{W}^*_{i+1,\ldots,m}(S^{(0)}), \\ 0 \end{Bmatrix}. \quad (16.17)$$

The control $-U^{(1)}, U^{(2)}$ or $U^{(3)}$, – at which is achieved this maximum, and there is conditional optimum control at i-th step/pitch $U_i^*(S^{(k)})$.

The optimization of the first step/pitch is implemented on the same principle, with the difference that initial state $S^{(0)}$ is not by chance and hypotheses about it to make not necessary.

If to the system, which is found in state $S^{(0)}$, on the first step/pitch control $U^{(1)}$, then it then passes into state $S^{(1)}$, and average/mean prize will be

$$\overline{W}_{1,2,\ldots,m}(S^{(0)},\ U^{(1)}) = Ap - s + \overline{W}^*_{2,3,\ldots,m}(S^{(1)}). \quad (16.18)$$

If will be used $U^{(2)}$, then

$$\overline{W}^*_{1,2,\ldots,m}(S^{(0)},\ U^{(2)}) =$$
$$= Ap - (s+r) + (1-p)\overline{W}^*_{2,3,\ldots,m}(S^{(0)}). \quad (16.19)$$

If will be used $U^{(3)}$, then

$$\overline{W}^*_{1,2,\ldots,m}(S^{(0)},\ U^{(3)}) = 0. \quad (16.20)$$

Page 169.

Optimum control on the first step/pitch is found from the condition

$$\overline{W}^{*}=\overline{W}^{*}_{1,2,\ldots,m}(S^{(0)})=$$
$$=\max\left\{\begin{array}{c}Ap-s+\overline{W}^{*}_{2,3,\ldots,m}(S^{(1)}),\\ Ap-(s+r)+(1-p)\overline{W}^{*}_{2,3,\ldots,m}(S^{(0)}),\\ 0\end{array}\right\}. \quad (16.21)$$

Thus, the construction of optimum control by random process is completed.

Let us consider a specific example. Let m=6; p=0.2; s=1; r=0.4; A=8. Let us determine optimum control.

Let us lead the optimization of latter/last (sixth) step/pitch. If after five steps/pitches shooting is already ended $(S^{(6)})$, then we apply control $U^{(3)}$ (we do not shoot) we obtain at the sixth step/pitch average/mean prize $\overline{W}^{(6)}(S^{(6)})=0$. If after five steps/pitches state $S^{(k)}(0<k<5)$, then the selection of optimum control is achieved/reached determined by the sign of the difference:

$$A(1-p)^{k}p-s; \qquad (16.22)$$

if it is more than zero, it is applied control $U^{(1)}$, if less than zero - control $U^{(3)}$. Let us compute the values of difference (16.22) for different values of k and let us reduce them to the table (see Table 16.1). In the same table in two lower lines let us give optimum control at sixth step/pitch $U^{*}_{6}(S^{(k)})$ and corresponding maximum average/mean prize $\overline{W}^{*}_{6}(S^{(k)})$. In the latter/last chair of the same table let us place the same data for case $S_{5}=S^{(6)}$.

Table 16.1.

k	0	1	2	3	4	5	6
$A(1-p)^k p - s$	0,60	0,28	0,02	<0	<0	<0	—
U_6^*	$U^{(1)}$	$U^{(1)}$	$U^{(1)}$	$U^{(3)}$	$U^{(3)}$	$U^{(3)}$	$U^{(3)}$
$\overline{W}_6^*(S^{(k)})$	0,60	0,28	0,02	0	0	0	0

Page 170.

Table 16.1 contains the full/total/complete results of the optimization of latter/last step/pitch. It is possible to formulate them as follows. If were produced already five shots and the latter/last information about the target (target is not affected) they acted recently (k=0) either for one shot to that (k=1), or for two shots (k=2), then it is necessary to produce latter/last shot, moreover without the exploration (control $U^{(1)}$). But if from the time of obtaining the latter/last information about the target are already released three or more than shots, then should be ceased shooting.

Thus, our optimum behavior at the sixth step/pitch in the accuracy is defined, no matter how unfolded the events at the previous five steps/pitches.

To optimize control at the fifth step/pitch we will be, using formula (16.13) and bearing in mind that function $\overline{W}_k^*(S^{(k)})$ with any k already is in Table 16.1. Here we had to compare not two, but three numbers, which correspond to three controls; however, since one of them corresponding to control $U^{(3)}$ — always zero, then write out we will be only two of them. If both numbers prove to be positive, let us select the control which corresponds greater of them; if one of the numbers will be positive, another - negative, let us select the control which corresponds to positive number; if both will be negative, let us select control $U^{(3)}$. Data let us reduce in Table 16.2.

The optimization of the first step/pitch is carried out. The output: after the fourth step/pitch under no conditions it is not necessary to send exploration.

Table 16.2.

k	0	1	2	3	4	6
$A(1-p)^k p - s + \overline{W}_6^*(S^{(k+1)})$	0,88	0,30	0,02	< 0	< 0	—
$A(1-p)^k p - (s+r) + (1-p)^{k+1} \overline{W}_5^*(S^{(0)})$	0,68	0,26	< 0	< 0	< 0	—
$U_5^*(S^{(k)})$	$U^{(1)}$	$U^{(1)}$	$U^{(1)}$	$U^{(3)}$	$U^{(3)}$	$U^{(3)}$
$\overline{W}_{5,6}^*(S^{(k)})$	0,88	0,30	0,02	0	0	0

Page 171.

If the information about the target is obtained recently or for one or two shots thus far, then it is necessary to do the fifth shot (control $U^{(1)}$). but if from the moment/torque of obtaining the latter/last information about the target are done three or more than shots, it is necessary to cease shooting.

Let us compose analogous table for the optimization of control at the fourth step/pitch (Table 16.3).

Optimization of the fourth step/pitch it is carried out. The output: if after the third step/pitch system is in state $S^{(0)}$ (are recently obtained the information about the target), then at the fourth step/pitch is equally profitable any of the controls $U^{(1)}$. $U^{(2)}$— to do a shot and to send or not to send exploration. If the information about the target is obtained for one shot or for two to that, optimum control will $U^{(2)}$ — do a shot and sand exploration. If the information about the target is obtained for three shots, optimum controls $U^{(3)}$ — to cease shooting.

The optimization of the third and second step/pitch is carried out in Tables 16.4 and 16.5.

During the optimization of the first step/pitch there is no necessity to vary the results of previous. The state of the system before the first step/pitch exists $S^{(0)}$. Therefore in the appropriate table there will be only one column, which corresponds $S^{(0)}$ (see Table 16.6).

The process of the optimization of control is completed. Is found the optimum control: unconditional - on the first step/pitch and conditional on all others:

$$U^{*}=(U_{1}^{*}=U^{(1)},\ U_{2}^{*}(S_{1}),\ U_{3}^{*}(S_{2}),\ U_{4}^{*}(S_{3}),\ U_{5}^{*}(S_{4}),\ U_{6}^{*}(S_{5})).$$

Table 16.3.

k	0	1	2	3	6
$A(1-p)^k p - s + \overline{W}^*_{5,6}(S^{(k+1)})$	0,90	0,30	0,02	<0	—
$A(1-p)^k p - (s+r) + {} + (1-p)^{k+1}\overline{W}^*_{5,6}(S^{(0)})$	0,90	0,34	0,07	<0	—
$U^*_4(S^{(k)})$	$U^{(1)}, U^{(2)}$	$U^{(2)}$	$U^{(2)}$	$U^{(3)}$	$U^{(3)}$
$\overline{W}^*_{4,5,6}(S^{(k)})$	0,90	0,34	0,07	0	0

Page 172.

The corresponding average/mean prize $\overline{W}^*=1.08$; in other words, the maximum average/mean income which we can attain, rationally combining shooting with the exploration, is equal to 1.08.

Table 16.4.

k	0	1	2	6
$A(1-p)^k p - s + \overline{W}^*_{4,5,6}(S^{(k+1)})$	0,94	0,35	0,02	—
$A(1-p)^k p - (s+r) + (1-p)^{k+1}\overline{W}^*_{4,5,6}(S^{(0)})$	0,92	0,46	0,08	—
$U_3^*(S^{(k)})$	$U^{(1)}$	$U^{(2)}$	$U^{(2)}$	$U^{(3)}$
$W^*_{3,4,5,6}(S^{(k)})$	0,94	0,46	0,08	0

Table 16.5.

k	0	1	6
$A(1-p)^k p - s + \overline{W}^*_{3,4,5,6}(S^{(k+1)})$	1,06	0,36	—
$A(1-p)^k p - (s+r) + (1-p)^{k+1}\overline{W}^*_{3,4,5,6}(S^{(0)})$	0,95	0,48	—
$U_2^*(S^{(k)})$	$U^{(1)}$	$U^{(2)}$	$U^{(3)}$
$\overline{W}^*_{2,3,4,5,6}(S^{(k)})$	1,06	0,48	0

Table 16.6.

k	0
$A(1-p)^k p - s + \overline{W}^*_{2,3,4,5,6}(S^{(k+1)})$	1,08
$A(1-p)^k p - (s+r) + (1-p)^{k+1}\overline{W}^*_{2,3,4,5,6}(S^{(0)})$	1,05
U_1^*	$U^{(1)}$
$\overline{W}^* = W^*_{1,2,3,4,5,6}$	1,08

Page 173.

Let us look how occurs the realization of optimum control.

At the first step/pitch we always apply control $U^{(1)}$, i.e. we make shot and we do not send exploration. As a result of this the system passes into state $S^{(1)}$. In this state, as can be seen from Table 16.5, optimum control is $U^{(2)}$ — to do a shot and to send exploration. If exploration communicates "target it is affected", then shooting ceases, i.e., at all subsequent steps/pitches is applied control $U^{(3)}$. But if exploration communicates "target it is not affected", then system passes into state $S^{(0)}$. In this state, as can be seen from Table 16.4, at the third step/pitch should be to apply control $U^{(1)}$ (done a shot and not sent the exploration), as a result of which the system will pass into state $S^{(1)}$.

Being converted to Table 16.3, we find optimum control on fourth step/pitch $U^{(2)}$ (to shoot and to send exploration). If will be reported "target it is affected", shooting ceases at the following step/pitch. If will be reported "target it is not affected", then system passes into state $S^{(0)}$: from Table 16.2 we find optimum control on fifth step/pitch $U^{(1)}$ (to shoot and not to send exploration). As a result of this control the system will pass into state $S^{(1)}$. From Table 16.1 it is evident that in this case the optimum control at the sixth step/pitch again exists $U^{(1)}$.

Fig. 16.3 depicts the trajectory of point S in the phase space during the obtained optimum control, constructed on the assumption that the first exploration communicated "target it was not affected", but the second - "target was affected".

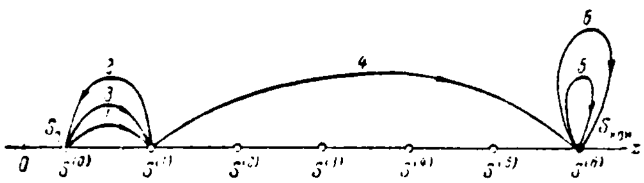


Fig. 16.3.

Page 174.

In our example it turned out that in the optimum control U* are included only controls $U^{(1)}$ and $U^{(2)}$, and control $U^{(3)}$ appears only in the case Proceedings "target it is affected". This will not always be thus. In other parameters of the problem control $U^{(3)}$ can prove to be advantageous and without obtaining Proceedings the "target is affected". We recommend to reader as the exercise to find the solution of problem with the following parameters:

$$m=5;\quad p=\frac{1}{2};\quad A=1;\quad s=\frac{1}{9};\quad r=\frac{1}{7}.$$

Let us give for the collation the optimum control which must be obtained during the solution of this task

$$U^{*}=(U^{(1)},\ U^{(1)},\ U^{(1)},\ U^{(3)},\ U^{(3)}).$$

This stochastic task proves to be in a certain sense "degenerate," since optimum control is not by chance.

REFERENCES

1. R. Bellman. Dynamic programming, IL, Moscow, 1960.

2. "Contemporary mathematics for engineers". coll. edited by E. F. Bekkenbakh, IL, 1958.

3. Ya. S. Venttsel'. Introduction to operations research. Publishing house "Soviet radio", 1964.

4. A. A. Krasovskiy, G. S. Pospelov. Bases of automation and technical cybernetics, Energoizdat, M. - L., 1962.

END

DATE
FILMED
5 81

DTIC