RR-74-8

ON OPTIMAL COMPROMISE FOR MULTIDIMENSIONAL RESOURCE DISTRIBUTION

Alexander M. Letov

Yuri A. Rozanov

June 1974

On Optimal Compromise for Multidimensional Resource Distribution

Alexander M. Letov and Yuri A. Rozanov

Introduction

Let us consider a resource $W > 0$ is to be divided in parts $w_1,\ldots,w_n$ where ($\sum_1^n w_n = W$) and consumed in the following way: at the first step, an amount x_1,

$$\alpha_1 \leq x_1 \leq w_1 - b_1 \quad ,$$

can be consumed; at the second step, one can consume x_2,

$$\alpha_2 \leq x_2 \leq y_1 + w_2 - b_2 \quad ,$$

where

$$y_1 = w_1 - x_1 \quad ,$$

and, generally, at the kth step, consumption is x_k,

$$a_k \leq x_k \leq y_{k-1} + w_k - b_k \quad , \tag{1}$$

where a_k, b_k are some non-negative constants and y_k,

$$y_0 = 0 \quad , \quad y_k = \sum_{j=1}^{k} w_j - \sum_{j=1}^{k} x_j \quad ; \quad k = 1,\ldots,n.$$

Of course, all parameters a_k, b_k and w_k $(k = 1,\ldots,n)$ are such that there is at least one feasible resource distribution, say

$$x_1 = a_1, \ldots, x_n = a_n \quad .$$

The problem is to find the "best" distribution of resources. If it is possible to assume that the distribution $x_1, \ldots, x_n$ of resources gives us a total benefit

$$u(x) = \sum_1^n c_k x_k \quad ,$$

where $c_1, \ldots, c_n$ are some coefficients, then the finding of the best feasible distribution of resources becomes a classical linear programming problem.

But suppose there are, in some sense, <u>independent</u> consumers who are not very much interested in the <u>total</u> benefit, and the desirable purpose of each of them is to receive as much as possible. Suppose these consumers (who are not living in a jungle!) want to reach a compromise based on some <u>reasonable</u> demands.

Namely, suppose it is given that some demands $x_1^*, \ldots, x_n^*$ are considered as quite reasonable by all consumers, yet these demands are <u>not feasible</u>: $x_1^*, \ldots, x_n^*$ do not satisfy the constraints (1). For example, the total demand $\sum_1^n x_k^*$ can be much more than the total resources amount $W = \sum_1^n w_k$. The problem is how to distribute our resources $w_1,\ldots,w_n$ according to the demands $x_1^*,\ldots,x_n^*$, which contradict each other in the sense that if we satisfy one group of consumers, then we leave too little for others.

A solution of this problem may be based on minimization (in the proper sense) of some "distance" between resource distribution vector $x = (x_1,\ldots,x_n)$ and the demand vector $x^* = (x_1^*,\ldots,x_n^*)$. This is considered below.

We wish to say that our problem on a compromise for many consumers which was described above arose in the water resource distribution field.

Example. One can realize a problem of water storage and of water distribution during some n sequential periods of time. Suppose during each kth period, the storage receives the water amount w_k, and some amount x_k, $x_k \geq a_k$, is taken in such a way that the rest of water resource will not be less than $b_k \geq 0$; $k = 1,\ldots,n$. If we are given the desirable demands $x_1^*,\ldots,x_n^*$ for the water from this storage (x_k^* for the kth period of time), then we have to deal with the problem described above.

Example. Let us consider a big river basin which is divided, according to geographic or economic principles, in n sequential parts (along the main river). Let the total available amount of water (in the proper scale) at the kth region of this basin be w_k. Suppose at every region one can consume a corresponding water amount x_k, $x_k \geq a_k$, such that the rest has to be not less than some $b_k \geq 0$. Obviously, if at the first (k-1) parts it consumed amounts $x_1, \ldots, x_{k-1}$, then x_n is bounded with the value $\sum_{j=1}^{k} w_j - \sum_{j=1}^{k-1} x_j - b_n$. Under the desirable but non-realistic water demands $x_1^*, \ldots, x_n^*$ for all n regions of the river basin, we have to deal with the problem on a compromise concerning the actual water distribution.

In the general situation of the resources shortage, when it is reasonable to assume that

$$x_k \leq x_k^* \quad ; \qquad k = 1, \ldots, n, \tag{2}$$

we suggest that to determine a distance between distribution vector $x = (x_1, \ldots, x_n)$ and demand vector $x^* = (x_1^*, \ldots, x_n^*)$ as

$$R(x, x^*) = \sqrt{\sum_{1}^{n} \lambda_k (x_k - x_k^*)^2} \quad , \tag{3}$$

where $\lambda_1, \ldots, \lambda_n$ are some non-negative coefficients. The choice of the proper $\lambda_1, \ldots, \lambda_n$ may be considered as re-evaluation of different demands $x_1^*, \ldots, x_n^*$ under certain circumstances. For example, some demands x_k^* may be neglected completely (under the choice of $\lambda_k = 0$). But we suppose the choice of weight-coefficients $\lambda_1, \ldots, \lambda_n$ is such that all consumers agree to consider the corresponding metric $r(x, x^*)$ as the loss function, i.e. a vector x' is preferable with respect to a vector x" if

$$r(x',x^*) \leq r(x'',x^*) \quad .$$

According to this agreement, the most preferable distribtution vector will be

$$x^o = (x_1^o,\ldots,x_n^o) \quad ,$$

for which

$$r(x^o,x^*) = \min_x r(x,x^*) \quad , \tag{4}$$

where x runs all possible distribution, i.e. $x = (x_1,\ldots,x_n)$ satisfies to the constraints (1) and (2).

Unfortunately, all components $x_1^o,\ldots,x_n^o$ of such minimum points generally depend on <u>all</u> resource components $w_1,\ldots,w_n$, and if we have to choose the amount x_k only with our knowledge of $w_1,\ldots,w_k$; $x_1,\ldots,x_{k-1}$, then we actually choose the proper x_k^o; $k = 1,\ldots,n$.

But sometimes we can assume that all $w_1,\ldots,w_n$ actually already are known at the first step. Say for a water storage or a basin with one big river, the components $w_2,\ldots,w_n$ may be much less than w_1, and in this case we can assume approximately that

$$w_1 = W \quad ; \qquad w_2 = \cdots = w_n = 0 \quad . \tag{5}$$

Let us consider the case when all resource components $w_1,\dots,w_n$ are known from the very beginning, and we can assume minimum point $x = (x_1^o,\dots,x_n^o)$ of the loss function $r(x,x^*)$ as the <u>optimal compromise</u> for our resource distribution problem.

The demands $x_1^*,\dots,x_n^*$, which generally are implicit functions of actual resources $w_1,\dots,w_n$, in this case are known constants, and the minimization problem for loss function $r(x,x^*)$ of the type (3) is a problem of quadratic programming. Namely, after a variables substitution,

$$w_1 - (a_1 + b_1) \to w_1$$

$$w_2 - (a_2 + b_2 - b_1) \to w_2$$

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

$$w_n - (a_n + b_n - b_{n-1}) \to w_n \quad ,$$

$$x_k - a_k \to x_k \quad ; \qquad k = 1,\dots,n,$$

the constraints (1) and (2) can be described as

$$0 \le x_k \le y_{k-1} + w_k \quad ; \qquad k = 1,\dots,n, \tag{1'}$$

where

$$y_0 = 0 \quad , \qquad y_k = \sum_{j=1}^{k} w_j - \sum_{j=1}^{k} x_j \quad ,$$

or in another form as

$$\sum_{j=1}^{k} x_j \leq \sum_{j=1}^{k} w_j \quad ; \qquad k = 1,\ldots,n. \tag{1''}$$

(Remember that $0 \leq x_k \leq x_k^*$, according to our assumption [2]).

1. Let us try to give some explicit formulas for the optimal $x_1^o,\ldots,x_n^o$, using the well-known Bellman's principle of dynamic programming. (See, for example, [2].)

Namely, let us begin with minimization of

$$\lambda_n (x_n - x_n^*)^2 \quad ,$$

where

$$0 \leq x_n \leq \min\,(y_{n-1} + w_n,\ x_n^*) \quad .$$

Obviously, x_n^o as a function of

$$\xi_n = y_{n-1} + w_n = \sum_{k=1}^{n} w_k = \sum_{k=1}^{n-1} x_k$$

is the following:

$$x_n = \begin{cases} \xi_n \ , & \xi_n \leq x_n^* \\ x_n^* \ , & \xi_n \geq x_n^* \end{cases} \quad .$$

(See Figure 1.)

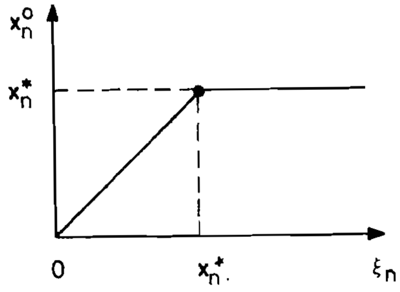


FIGURE 1

Remember that

$$y_{n-1} = \sum_{k=1}^{n-1} w_k - \sum_{k=1}^{n-1} x_k \quad .$$

Let us fix $x_1,\ldots,x_{n-2}$ and set

$$z_{n-1} = \sum_{k=1}^{n-1} w_k - \sum_{k=1}^{n-2} x_k + w_n \quad .$$

Then we have

$$x_n^o = \min\,(z_{n-1} - x_{n-1},\, x_n^*) \quad .$$

If $z_{n-1} > x_n^*$, then there is a feasible amount $x_{n-1} \geq 0$ such that $z_{n-1} - x_{n-1} \geq x_n^*$ and $x_n^o = x_n^*$. It holds true, for example, in the case of $w_n \geq x_n^*$, when the optimal x_{n-1}^o is obviously similar to x_n^o, namely,

$$x_{n-1}^o = \max\,(y_{n-2} + w_{n-1},\, x_{n-1}^*) \quad .$$

Suppose $\underline{w_n < x_n^*}$. If $z_{n-1} \geq x_n^* + x_{n-1}^*$, then

$$y_{n-2} + w_{n-1} > x^*_{n-1} \quad , \quad y_{n-1} + w_{n-1} - x^*_{n-1} + w_n > x^*_n \quad ,$$

and not only the last demand will be satisfied ($x^o_n = x^*_n$) but also $x^o_{n-1} = x^*_{n-1}$. In the case

$$0 \leq c_{n-1} = z_{n-1} - x^*_n < x^*_{n-1} \quad ,$$

we obviously have

$$\lambda_{n-1} (x_{n-1} - x^*_{n-1})^2 + \lambda_n (x^o_n - x^*_n)^2$$

$$= \begin{cases} \lambda_{n-1} (x_{n-1} - x^*_{n-1} - x^*_{n-1})^2 \quad , & \text{if } 0 \leq x_{n-1} \leq c_{n-1} \\ \lambda_{n-1} (x_{n-1} - x^*_{n-1})^2 + \lambda_n (c_{n-1} - x_{n-1})^2 \quad , & \\ & \text{if } c_{n-1} \leq x_{n-1} \leq x^*_{n-1} \end{cases}$$

and for small $\Delta x_{n-1} \geq 0$ the corresponding increment at the point $x_{n-1} = c_{n-1}$ of our utility function (under fixed $x_1,\ldots,x_{n-2}$) is

$$2\lambda_{n-1} (c_{n-1} - x^*_{n-1}) \Delta x + 0 (\Delta x^2) < 0 \quad ,$$

so the minimum point x^o_{n-1} is such that $x^o_{n-1} > c_{n-1}$. The same argument concerning $\Delta x_{n-1} < 0$ and $x_{n-1} = x^*_{n-1}$ gives us the inequality $x^o_{n-1} < x^*_{n-1}$. It means that if we cannot satisfy <u>both</u> of the demands x^*_n and x^*_{n-1}, then under the optimal compromise we have the strict inequalities

$$x^o_{n-1} < x^*_{n-1} \quad , \qquad x^o_n < x^*_n \quad .$$

In this case the function

$$f_{n-1}\,(x_{n-1}) = \lambda_{n-1}\,(x_{n-1} - x^*_{n-1})^2 + \lambda_n\,(x^o_n - x^*_n)^2$$

$$= \lambda_{n-1}\,(x_{n-1} - x^*_{n-1})^2 + \lambda_n\,(c_{n-1} - x_{n-1})^2 \quad ,$$

where

$$c_{n-1} = \sum_{k=1}^{n} w_k - \sum_{k=1}^{n-2} x_k - x^*_n$$

generally is not necessarily positive, but has the <u>absolute minimum point</u>

$$\tilde{x}_{n-1} = \frac{\lambda_{n-1} x^*_{n-1} + \lambda_n c_{n-1}}{\lambda_{n-1} + \lambda_n} \quad .$$

Because the considered function is monotone descreasing for $x_{n-1} \leq \tilde{x}_{n-1}$ and it is monotone increasing for $x_{n-1} \geq \tilde{x}_{n-1}$, we obviously obtain

$$x^o_{n-1} = \begin{cases} 0 \quad , & \tilde{x}_{n-1} \leq 0 \\ \tilde{x}_{n-1} \quad , & 0 \leq \tilde{x}_{n-1} \leq y_{n-2} + w_{n-1} \\ y_{n-2} + w_{n-1} \quad , & \tilde{x}_{n-1} \geq y_{n-2} + w_{n-1} \end{cases} \quad .$$

(Remember, as it was shown above, that under the conditions

$c_{n-1} < x^*_{n-1}$, $w_n < x^*_n$ the optimal amount x^o_{n-1} is strictly less than x^*_{n-1}.)

Let us consider x^o_{n-1} as a function of

$$\xi_{n-1} = y_{n-2} + w_{n-1} = \sum_{k=1}^{n-1} w_k - \sum_{k=1}^{n-2} x_k \quad ,$$

which is the total amount of available resources at the (n-1)th step. If the next demand x^*_n is comparatively high and ξ_{n-1} is too small, namely,

$$0 \le (x^*_n - w_n) - \frac{\lambda_{n-1}}{\lambda_n} x^*_{n-1}$$

and

$$\xi_{n-1} \le (x^*_n - w_n) - \frac{\lambda_{n-1}}{\lambda_n} x^*_{n-1} \quad ,$$

or that is the same as

$$\tilde{x}_{n-1} = \frac{\lambda_{n-1} x^*_{n-1} + \lambda_n (\xi_{n-1} + w_n - x^*_n)}{\lambda_{n-1} + \lambda_n} \le 0 \quad ;$$

then $x^o_{n-1} = 0$. It is easy to verify that the first inequality always implies

$$x^*_{n-1} - \frac{\lambda_n}{\lambda_{n-1}} (x^*_n - w_n) \le 0 \quad , \qquad \xi \ge \tilde{x}_{n-1} \quad .$$

The function x^o_{n-1} of ξ_{n-1} may be one of the following types

$$\text{(a)} \qquad x^{o}_{n-1} = \begin{cases} 0\ , & 0 \le \xi_{n-1} \le (x^*_n - w_n) - \dfrac{\lambda_{n-1}}{\lambda_n}\, x^*_{n-1} \\ \tilde{x}_{n-1}\ , & x^*_n - w_n - \dfrac{\lambda_{n-1}}{\lambda_n}\, x^*_{n-1} \le \xi \\ & \le x^*_{n-1} + x^*_n - w_n \\ x^*_{n-1}\ , & \xi_{n-1} \ge x^*_{n-1} + x^*_n - w_n \end{cases} \tag{6}$$

or

$$\text{(b)} \qquad x^{o}_{n-1} = \begin{cases} \xi\ , & 0 \le \xi_{n-1} \le x^*_{n-1} - \dfrac{\lambda_n}{\lambda_{n-1}}\,(x^*_n - w_n) \\ \tilde{x}_{n-1}\ , & x^*_{n-1} - \dfrac{\lambda_n}{\lambda_{n-1}}\,(x^*_n - w_n) \le \xi_{n-1} \\ & \le x^*_{n-1} + x^*_n - w_n \\ x^*_{n-1}\ , & \xi_{n-1} \ge x^*_{n-1} + x^*_n - w_n\ . \end{cases} \tag{7}$$

(See Figure 2.)

In particular, we obtain that under the condition of a non-extreme resources shortage, when we don't use extreme strategy 0 or ξ_{n-1} ("nothing" or "all"), the optimal amount x^{o}_{n-1} coincides with $\tilde{x}_{n-1}$, which is the <u>linear function</u> of ξ_{n-1} as well as of the parameters w_n; x^*_n, x^*_{n-1}, namely,

$$x^{o}_{n-1} = \tilde{x}_{n-1} = \frac{\lambda_{n-1} x^*_{n-1} + \lambda_n\,(\xi_{n-1} + w_n - x^*_n)}{\lambda_{n-1} + \lambda_n}\ , \tag{8}$$

where

$$\xi_{n-1} = y_{n-2} + w_{n-1} = \sum_{k=1}^{n-1} w_k - \sum_{k=1}^{n-2} x_k\ .$$

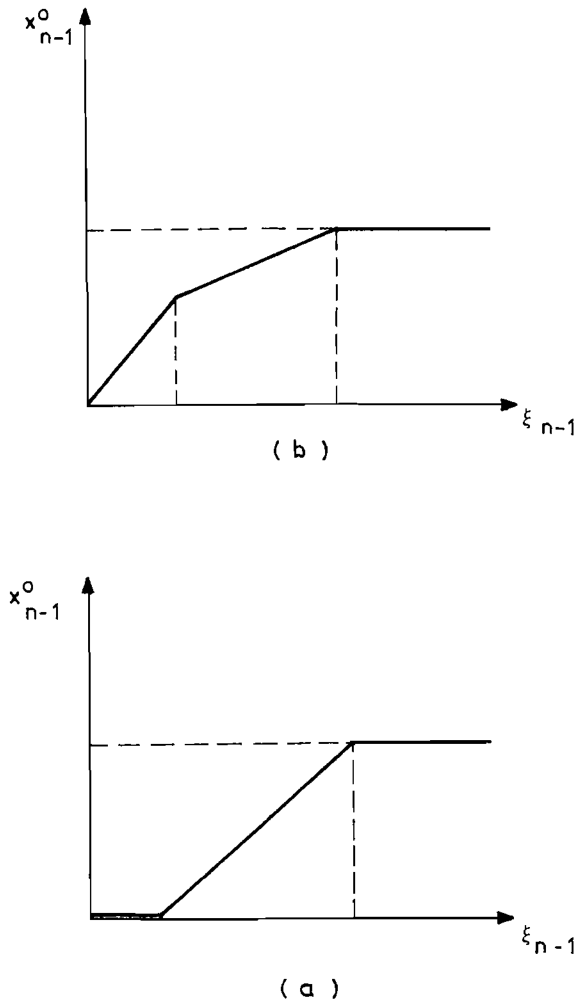


FIGURE 2

In a very similar way one can consider the structure of all other optimal components x_k^o; $k = n-2, n-1, \ldots, 1$, which are the minimum points of the corresponding functions

$$f_k(x_k) = \lambda_k (x_n - x_k^*)^2 + \sum_{j=k+1}^{n} \lambda_j (x_j^o - x_j^*)^2 \quad , \tag{9}$$

$$0 \leq x_k \leq \min(\xi_k, x_k^*) \quad ,$$

where

$$\xi_k = y_{k-1} + w_k = \sum_{i=1}^{k} w_i - \sum_{i=1}^{k-1} x_i$$

is the total resources (available at the <u>kth</u> step).

In particular, it is very easy to discover the following properties of optimal distributions.

If $x_k^o = x_k^*$, then $x_j^o = x_j^*$ for all $j > k$; moreover, if $x_k^o \neq 0$,

$$\lambda_k (x_k^* - x_k^o) \geq \max_{k<j\leq n} \lambda_j (x_j^* - x_j^o) \quad . \tag{10}$$

Indeed in the contrary case,

$$\lambda_k (x_k^* - x_k^o) < \lambda_i (x_i^* - x_i^o)$$

for some $i > k$, and if we take

$$x_k = x_k^o - \Delta x \quad , \quad x_i = x_i^o + \Delta x \quad , \quad x_j = x_j^o \ (j \neq i, k) \quad ,$$

then we could improve the optimal distribution, because

$$-\lambda_k \; (x_k^o - x_k^*)^2 + \lambda_i \; (x_i^o - x_i^*)^2 + \lambda_k \; (x_k^o - \Delta x - x_k^*)$$

$$+ \lambda_i \; (x_i^o + \Delta x - x_i^*)^2$$

$$= 2 \; [\lambda_k \; (x_k^* - x_k^o) - \lambda_i \; (x_i^* - x_i^o)] \; \Delta x + 0 \; (\Delta \tilde{x}) < 0$$

for sufficiently small $\Delta x > 0$.

Suppose that

$$x_k^o < x_k^* \quad ; \qquad k = 1, \ldots, n,$$

what one may expect normally under the resources shortage, then at each step $k = 1, \ldots, n$ there are three possibilities:

$$1) \quad x_k = 0 \qquad 2) \quad 0 < x_k < \xi_k \qquad 3) \quad x_k = \xi_n \quad .$$

(Remember that ξ_k is the total resources which are available at the <u>kth</u> step.)

According to the inequality (10), the decision $x_k^o = \xi_n$ may be optimal only if there is a possibility to satisfy other demands $x_{k+1}^*, \ldots, x_n^*$ (with their own resources $w_{k+1}, \ldots, w_n$) in such a way that

$$\max_{k<j\leq n} \lambda_j \; (x_j^* - x_j^o) \leq \lambda_k \; (x_k^* - \xi_k) \quad .$$

So we can check whether or not it is possible that $x_k^o = \xi_k$; $k = 1,2,\dots,n$. For example, in the case when

$$w_1 = W \quad ; \qquad w_2 \approx \cdots \approx w_n = 0$$

and the demands x_k^*; $k = 1,\dots,n$ don't decrease very much, we have $x_k^o < \xi_k$, $k = 1,2,\dots,n$.

On the other hand, it usually is not worth transferring a large amount $y_n = \xi_k - x_n$ to other consumers when the demand x_n^* is large. In particular, if $y_k^o > 0$ and

$$\lambda_k x_k^* > \max_{k<j\leq n} \lambda_j \, (x_j^* - x_j) \tag{11}$$

for some distribution $(x_{k+1},\dots,x_n)$ of the resources $y_k^o + w_{k+1},\dots,w_n$, then $x_n^o \neq 0$, because in the case $x_k^o = 0$ with taking back a sufficiently small amount Δy_k we can decrease our utility function $r(x,x^*)$.

Thus, one can sometimes find out (without any calculations) that the optimal distribution has to be the following:

$$0 < x_k^o < \xi_k \quad ; \qquad k = 1,2,\dots,n, \tag{12}$$

i.e. at each step k, one consumes something but not all available amount ξ_k.

In this case the minimum points x_n^o of the corresponding functions $f_k(x_k)$--see (9)--can be determined in an obvious way. Namely,

$$x_n^o = \xi_n \quad , \qquad x_{n-1}^o = \tilde{x}_{n-1}$$

is the absolute minimum point of the parabolic function $f_{n-1}(x_{n-1})$--see (8). The next function $f_{n-2}(x_{n-2})$ also is of the same parabolic type because $x_{n-1}^o = \tilde{x}_{n-1}$ is a linear function of $\xi_{n-1} = \sum_{k=1}^{n-1} w_k - \sum_{k=1}^{n-2} x_k$ (as well as of the parameters w_n, x_n^*, x_{n-1}^*), and under the condition (12): $0 < x_{n-1}^o < \xi_{n-2}$, the optimum x_{n-2}^o has to coincide with the absolute minimum point of the function $f_{n-2}(x_{n-2})$, and so on. By the same arguments, the optimum x_k^o for all other $K = n-2,\ldots,1$ coincides (under the condition [12]) with the absolute minimum point of the corresponding parabolic function

$$f_k(x_k) = \lambda_k (x_k - x_k^*)^2 + \sum_{j=k+1}^{n} (x_i^o - x_i^*)^2 \quad ,$$

where x_j^o, $j > k$ are the proper linear functions of $\xi_k = \sum_{i=1}^{k} w_i - \sum_{i=1}^{k-1} x_i$ as well as of the parameters $w_n,\ldots,w_{k+1}$; $x_n^*,\ldots,x_k^*$.

Remember that under the condition $x_n^o < x_n^*$ there are only three possibilities at each kth step: 1) $x_k^o = 0$ 2) $0 < x_k^o < \xi_k$ 3) $x_k^o = \xi_k$, where $\xi_k = \sum_{i=1}^{k} w_i - \sum_{i=1}^{k-1} x_i$ is the total available resource. So we can describe the type of our decision as the corresponding sequence $(d_1, d_2,\ldots,d_n)$, where d_k means "nothing"

($x_k^o = 0$), or "something" ($0 < x_k^o < \xi_k$), or "all" ($x_k^o = \xi_k$). Obviously if we know the type of the optimal distribution, then the optimal components x_k^o can be determined as $x_k^o = 0$, or as $x_k^o = \xi_k$, or as the absolute minimum point $x_k^o = \tilde{x}_k$ of the proper parabolic function $f_k(x_k)$--see (9)--with the already chosen $x_1^o,\dots,x_{k-1}^o$, which are the proper <u>linear functions</u> of $\xi_k = \sum_{i=1}^{k} w_i - \sum_{i=1}^{k-1} x_i$, as well as of the parameters $w_n,\dots,w_{k+1}$; $x_n^*,\dots,x_k^*$.

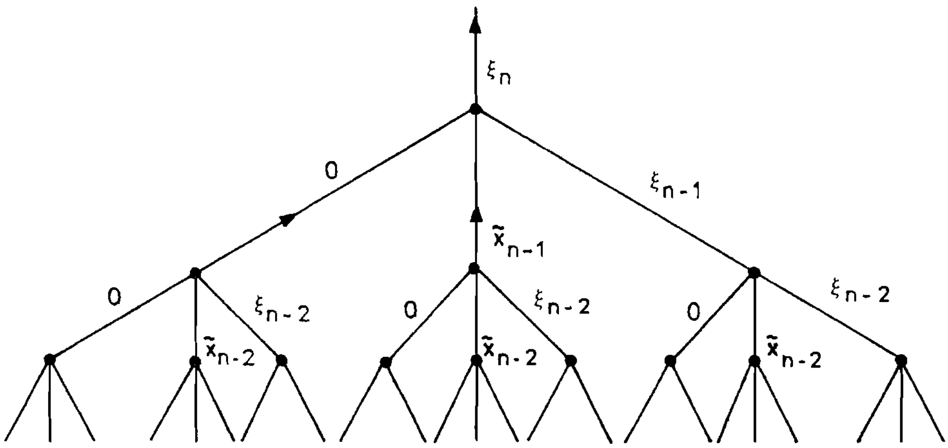


FIGURE 3

The tree of the possible decision types under the resources shortage when $x_k^o < x_k^*$ for all $k = 1,\dots,n$.

Example. For two consumers and resources $w_1 = W$, $w_2 = 0$ under the maximum demands $x_1^* = W$, $x_2^* = W$ for

$$r(x_1 x^*) = \sqrt{(x_1 - x_1^*)^2 + (x_2 - x_2^*)^2} \quad ,$$

we have

$$x_2^o = W - x_1 \quad , \qquad \tilde{x}_1 = \frac{W}{2} \quad .$$

Remember that $\tilde{x}_1$ is the absolute minimum point of the corresponding function

$$(x_1 - W)^2 + (x_2^o - W)^2 = (x_1 - W)^2 + x_1^2 \quad .$$

Because the value $\tilde{x}_1 = W/2$, it satisfies the constraints (1) and (2):

$$0 \leq \tilde{x}_1 \leq W \quad .$$

We obtain $x_1^o = \tilde{x}_1$ so our optimization principle gives us

$$x_1^o = W/2 \quad , \qquad x_2^o = W/2 \quad .$$

2. As it was described above, the optimal decision (even at the first step!) depends very much on all parameters $w_1,\ldots,w_n$. So a new problem arises in the case when the corresponding decision about a proper amount x_k has to be made with knowing only about $w_1,\ldots,w_k$ and $x_1,\ldots,x_{k-1}$.

Generally, in order to choose the components x_k^o (when $x_1^o,\dots,x_{k-1}^o$ have been chosen already and $w_1,\dots,w_{k-1}$ are known), it may be recommended to substitute the unknown parameters $w_{k+1},\dots,w_n$ with the appropriate estimate $w_{k+1}^*,\dots,w_n^*$ (which can be improved at the next (k+1)th step when x_k^o will be chosen and w_{k+1} will be known).

It is possible, for example, to use upper and lower boundaries for unknown resources. Namely, if we have some estimates

$$\underline{w}_k \leq w_k \leq \overline{w}_k \quad ; \qquad k = 1,\dots,n, \tag{13}$$

then we can obtain it as it was described above corresponding to optimal distribution vectors

$$\underline{x} = (\underline{x}_1,\dots,\underline{x}_n) \quad \text{and} \quad \overline{x} = (\overline{x}_1,\dots,\overline{x}_n)$$

with respect to the parameters

$$\underline{w} = (\underline{w}_1,\dots,\underline{w}_n) \quad \text{and} \quad \overline{w} = (\overline{w}_1,\dots,\overline{w}_n)$$

under the same demands $x_1^*,\dots,x_n^*$. By intuition it seems that

$$\underline{x}_k \leq x_k^o \leq \overline{x}_k \quad ; \qquad k = 1,\dots,n,$$

and actually it is true.

Let us show that for any parameters $\underline{w}_k \leq \overline{w}_k$; $k = 1,\dots,n$ the minimum points $\underline{x} = (\overline{x}_1,\dots,\overline{x}_n)$ and $\overline{x} = (\overline{x}_1,\dots,\overline{x}_n)$ of the loss function

$$r\ (x_1 x^*) = \sqrt{\sum_1^n \lambda_k\ (x_k - x_k^*)^2}$$

under the corresponding constraints (1) and (2) satisfy the inequalities $\underline{x}_k \leq \overline{x}_k$; $k = 1,\dots,n$.

Obviously, under the resources shortage, more precisely under the condition (2), the total consumption in the case of optimal distribution has to be as much as is available:

$$\sum_1^n x_k = \max\ .$$

So for the optimal distributions $\underline{x} = (\underline{x},\dots,\underline{x}_n)$ and $\overline{x} = (\overline{x}_1,\dots,\overline{x}_n)$ with respect to the resources $w = (w_1,\dots,w_n)$ and $\overline{w} = (\overline{w}_1,\dots,\overline{w}_n)$, we have

$$\sum_1^n \underline{x}_k \leq \sum_1^n \overline{x}_k\ .$$

If $\overline{x}_i < \underline{x}_i$ and $\Delta_i = \underline{x}_i - \overline{x}_i > 0$ for some i (say $i \varepsilon I$), then under $x_i = \overline{x}_i$ we have an extra positive amount $\Delta = \sum_{i \varepsilon I} \Delta_i$ (in comparison with $x_i = \underline{x}_i$, $i \varepsilon I$) which can be distributed among other variables x_j, $j \notin I$ (under the same resources $\underline{w}_1,\dots,\underline{w}_n$). Apparently, for the $\overline{x}_j \geq \underline{x}_j$ ($j \notin I$) we have

$$\sum_{j\varepsilon I} \overline{x}_j \geq \sum_{j\varepsilon I} \underline{x}_j + \Delta \ ,$$

so there is a partition $\Delta = \sum_{j \notin I} \Delta_j$ (with $\Delta_j \geq 0$) such that

$$x_j = \overline{x}_j - \Delta_j \geq \underline{x}_j \quad ; \quad j \notin I \ .$$

The distribution of the components

$$x_j = \overline{x}_i = \underline{x}_i - \Delta_i \ , \quad i\varepsilon I \ ; \quad x_j = \underline{x}_j + \Delta_j \ , \quad j \notin I$$

is feasible with $(x_j \leq \overline{x}_j !)$ under the resources $\underline{w}_1, \ldots, \underline{w}_n$. Because $\underline{x} = (\underline{x}_1, \ldots, \underline{x}_n)$ is the corresponding minimum point, we have

$$\sum_{i\varepsilon I} \lambda_i \ (x_i - \Delta_i - x_i^*) - \sum_{i\varepsilon I} \lambda_i \ (\underline{x}_i - x_i^*)^2$$

$$\geq \sum_{j \notin I} \lambda_j \ (\underline{x}_j - x_j^*)^2 - \sum_{j \notin I} \lambda_j \ (\underline{x}_j + \Delta_j - x_j^*)^2$$

$$> \sum_{j \notin I} \lambda_j \ (x_j - x_j^*)^2 - \sum_{j \notin I} \lambda_j \ (x_j + \Delta_j - x_j^*)^2 \ ,$$

because $x_j \geq \underline{x}_j$, $j \notin I$ and

$$(x_j - x_j^*)^2 - (x_j + \Delta_j - x_j^*)^2 \leq (\underline{x}_j - x_j^*)^2$$

$$- (\underline{x}_j + \Delta_j - x_j^*)^2$$

for any $\Delta_j \geq 0$; $x_j + \Delta_j \leq x_j^*$. (See Figure 4.) Thus for

$$\bar{x}_i = \underline{x}_i - \Delta_i \ (i \varepsilon I) \quad , \qquad \bar{x}_j = x_j + \Delta_j \ (j \notin I)$$

we obtain that

$$r\ (\bar{x}, x^*) = \sum_{i \varepsilon I} \lambda_i\ (\bar{x}_i - x_i^*)^2 + \sum_{j \notin I} \lambda_j\ (\bar{x}_j - x_j^*)^2$$

$$> \sum_{i \varepsilon I} \lambda_i\ (\underline{x}_i - x_i^*)^2 + \sum_{j \notin I} \lambda_j\ (\bar{x}_j - \Delta_j - x_j^*)^2$$

$$= r\ (x', x^*)$$

for feasible distribution (under the resources $\bar{w}_1,\ldots,\bar{w}_n$) $x' = (x_1',\ldots,x_n')$ with components

$$x_i' = \underline{x}_i > \bar{x}_i \quad , \qquad i \varepsilon I \quad ; \qquad x_j' = \bar{x}_j - \Delta_j \quad , \qquad i \varepsilon I \quad .$$

But it contradicts the fact that $\bar{x} = (\bar{x}_1,\ldots,\bar{x}_n)$ is the minimum point of $r(x, x^*)$ with respect to the resources $\bar{w}_1,\ldots,\bar{w}_n$, so our assumption on the strict inequalities $\bar{x}_i < \underline{x}_i$, $i \varepsilon I$ is not true and $\underline{x}_k \leq \bar{x}_k$ for all $k = 1,\ldots,n$.

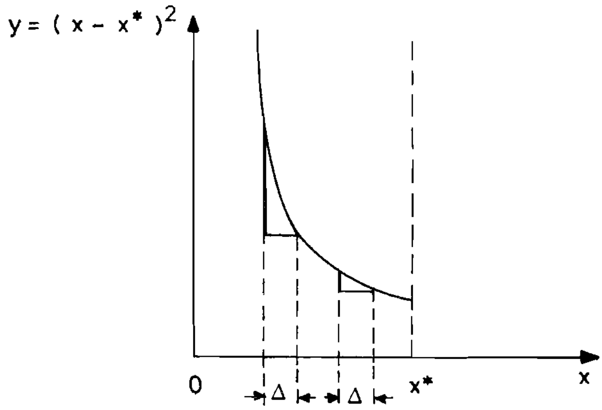


FIGURE 4

With knowledge of the inequalities, one can use the following resource distribution strategy in the case when at each kth step it is necessary to keep a good part of the total resources for the other consumers $k + 1,\dots,n$. Namely, one can choose the current amount x_k^o as the first component of the minimum point $(x_k^o,\dots,x_n^o)$ of the function

$$\sum_{j=k}^{n} \lambda_k \, (x_k - x_k^*)^2$$

under the constraints (1) and (2) with the already chosen $x_1^o,\dots,x_{k-1}^o$ and the corresponding parameters $w_j = \underline{w}_j$, $j = k + 1,\dots,n$, which are actually lower boundaries for the real resources. On the contrary, if one doesn't like to take great interest in other consumers, it is possible to use in a similar way the upper boundaries $w_j = \bar{w}_j^{(k)}$; $j = k+1,\ldots,n$.

Let us now consider $w_j,\ldots,w_n$ as random variables. In this case one can try to minimize a mean value of the loss function $r(x,x^*)$ and to find $\bar{x}^o = (\bar{x}_1^o,\ldots,\bar{x}_n^o)$ such that

$$Er\ (\bar{x}^{o}, x^*) = \min Er\ (x,x^*) \quad .$$

(Remember that each component x_k is a function of some data Γ_k including $x_1,\ldots,x_{k-1}$ and $w_1,\ldots,w_k$.)

Apparently the optimal decision at the last nth step is the same as it was above:

$$\bar{x}_n^o = x_n^o = \xi_n = \sum_{k=1}^{n} w_k - \sum_{k=1}^{n-1} x_k \quad .$$

According to the well-known Bellman's principle of dynamic programming, let us minimize the conditional expectation

$$f_{n-1}\ (x_{n-1}) = E\ \{\lambda_{n-1}\ (x_{n-1} - x_{n-1}^*)^2 + \lambda_n\ (\bar{x}_n^o - x_n^*)^2/\Gamma_{n-1}\}$$

$$= \lambda_{n-1}\ (x_{n-1} - x_{n-1}^*)^2 + \lambda_n\ (\eta_n - x_n^*)^2$$

$$+ \lambda_n\ E\ \{(\bar{x}_n^o - \eta_n)^2/\Gamma_{n-1}\} \quad ,$$

where

$$\eta_n = E\ \{\bar{x}_n^o/\Gamma_{n-1}\} \quad .$$

Obviously, $\bar{x}_n^o - \eta_n$ is the function only of w_n

$$(\bar{x}_n^o - \eta_n = w_n - E\{w_n/\Gamma_{n-1}\}) \quad ,$$

so η_n coincides with the optimal distribution component $x_n^o(w_1,\ldots,w_{n-1}, \bar{w}_n)$ with respect to the resources $w_1,\ldots,w_{n-1}$, $\bar{w}_{n-1} = E\{w_n/\Gamma_{n-1}\}$, and the same conclusion we have to make concerning $\bar{x}_{n-1}^o$, namely, $\bar{x}_{n-1}^o = x_{n-1}^o(w_1,\ldots,w_{n-1}, \bar{w}_n)$ is the optimal distribution component with respect to the parameters $w_1,\ldots,w_{n-1}$, w_n, where

$$\bar{w}_n - E\{w_n/\Gamma_{n-1}\} \quad .$$

Generally, $x_{n-1}^o(w_1,\ldots,w_{n-1}, \bar{w}_n)$ is the non-linear function of $\xi_{n-1} = \sum_{k=1}^{n-1} w_k - \sum_{k=1}^{n-2} x_k$ as well as of $\bar{w}_n$. See formulas (6) and (7).

Suppose that for all <u>possible parameters</u> $w = (w_1,\ldots,w_n)$, which may be from some known set W in n-dimensional vector space R^n, the corresponding optimal distributions $x^o = (x_1^o,\ldots,x_n^o)$ have to be of the <u>same type</u>. (It may be any of $\mathcal{N} = 3^{n-1}$ different types. For example, it may be of the type [12].) In this case $\bar{x}_n^o$ and $\bar{x}_{n-1}^o$ are known <u>linear functions</u> of the variables ξ_n and ξ_{n-1}, $\bar{w}_n$. Obviously, the conditional expectations

$$\eta_n = E\{\bar{x}_n^o/\Gamma_{n-2}\} \quad \text{and} \quad \eta_{n-1} = E\{\bar{x}_{n-1}^o/\Gamma_{n-2}\}$$

are of similar types and coincide with the corresponding optimal distribution components

$$x_n^o = x_n^o\ (w_1,\ldots,w_{n-2},\ \bar{w}_{n-1},\ \bar{w}_n)$$

and

$$x_{n-1}^o = x_{n-1}^o\ (w_1,\ldots,w_{n-2},\ \bar{w}_{n-1})$$

with respect to the resources $w_1,\ldots,w_{n-2}$, where

$$\bar{w}_{n-1} = E\ \{w_{n-1}/\Gamma_{n-2}\}\quad ,\qquad \bar{w}_n = E\ \{w_n/\Gamma_{n-2}\}\quad .$$

We have

$$f_{n-2}\ (x_{n-2}) = E\ \{\lambda_{n-2}\ (x_{n-2} - x_{n-2}^*)^2 + \sum_{j=n-1}^{n} \lambda_j\ (\bar{x}_j^o - x_n^*)^2\}$$

$$= \lambda_{n-2}\ (x_{n-2} + x_{n-2}^*)^2 + \sum_{j=n-1}^{n} \lambda_j\ (\eta_j - x_j^*)^2$$

$$+ \sum_{j=n-1}^{n} \lambda_j\ E\ \{(\bar{x}_j^o - \eta_j)^2/\Gamma_{n-2}\}\quad .$$

Here the last term is the <u>constant</u> (because $\bar{x}_j^o$ is the proper linear function, in particular, the linear functions of the variable x_{n-2}), so $\eta_j = x_j^o\ (w_1,\ldots,w_{n-2},\ w_{n-1},\ \bar{w})$ are the optimal distribution components with respect to the resources $w_1,\ldots,w_{n-2}$, $\bar{w}_{n-1}$, $\bar{w}_n$, and $\bar{x}_{n-2}^o$ also has to be the optimal component with respect to parameters $w_1,\ldots,w_{n-2}$, $\bar{w}_{n-1}$, $\bar{w}_n$.

Now it seems clear that at each kth step the optimal amount $\bar{x}_k^o$ as the minimum point of the function

$$f_k(x_k) = E\{\lambda_k (x_k - x_k^*)^2 + \sum_{j=k+1}^{n} \lambda_j (\bar{x}_j^o - x_j^*)^2 / \Gamma_k\}$$

coincides with the optimal distribution component with respect to the parameters $w_1, \ldots, w_k, \bar{w}_{k+1}, \ldots, \bar{w}_n$, where

$$\bar{w}_j = E\{w_j / \Gamma_k\} \quad ; \quad j = k+1, \ldots, n, \tag{14}$$

namely,

$$\bar{x}_k^o = x_k^o (w_1, \ldots, w_k, \bar{w}_{k+1}, \ldots, \bar{w}_n) \quad ; \quad k = 1, \ldots, n. \tag{15}$$

(Remember we assumed above that for all possible parameters $w_1, \ldots, w_n$ the corresponding optimal distributions are the same type!)

Note that in the case when, for different groups of parameters $(w_1, \ldots, w_n)$ the corresponding optimal distributions $(x_1^o, \ldots, x_n^o)$ are of different types and $\bar{x}_j^o$ are non-linear functions of the variables w_k, $k < j$, the optimal distributions $(x_1^o, \ldots, \bar{x}_n^o)$--concerning mean value of the loss function--is more complicated than it was described above.

It is worthy to note also that the mean value criterion is not uniformly good for any probability distributions of the parameters $(w_1, \ldots, w_n)$.

Example. Let us consider two demands x_1^*, x_2^* for resources with independent components w_1, w_2. Suppose $w_2 \neq 0$ with a very small probability p (say, $p = 0.001$), so we almost can be sure that $w_2 = 0$.

In order to make this more clear, let us assume that $E\,w_2 \geq x_2^*$. Then on the basis of mean value criterion we have to choose $\bar{x}_1^o = w_1$ and $\bar{x}_2^o = 0$ (with a big probability I-p). Obviously, such a decision is not good in the case when with a good guarantee the second demand x_2^* has to be partly satisfied.

3. Suppose that under a condition of unknown resources $w_1,\ldots,w_n$ the corresponding demands $x_1^*,\ldots,x_n^*$ are given in the form

$$x_n^* = \sum_{j=1}^{k} \alpha_{kj}^* w_j \quad ; \qquad x = 1,\ldots,n,$$

where α_{nj}^*, $0 \leq \alpha_{nj}^* \leq 1$, are some coefficients (i.e. the k<u>th</u> consumer demands α_{kj}^*<u>th</u> part of the resource w_j, $j \leq k$). Of course, these coefficients α_{kj}^* may be <u>non-feasible</u>. Namely, it may be $\sum_{k=j}^{n} \alpha_{nj}^* > 1$, and the problem is to find <u>feasible</u> coefficients α_{nj}:

$$\alpha_{nj} \geq 0 \quad , \qquad \sum_{k=j}^{n} \alpha_{nj} \leq 1 \qquad\qquad (16)$$

which are <u>optimal</u> in some reasonable sense for the resource distribution

$$x_n = \sum_{j=1}^{k} \alpha_{nj} w_j \quad ; \qquad k = 1,\ldots,n.$$

Note that the conditions (16) for <u>arbitrary</u> parameters $w_1,\ldots,w_n$ are equivalent to the following:

$$\sum_{k=1}^{m} x_n = \sum_{k=1}^{m} \left(\sum_{j=1}^{k} \alpha_{nj} w_j \right)$$

$$= \sum_{j=1}^{m} \left(\sum_{k=j}^{m} \alpha_{kj} \right) w_j \leq \sum_{j=1}^{m} w_j \quad ; \qquad m = 1,\ldots,n.$$

According to our general principle of optimality, we propose as an <u>optimal compromise</u> the resource distribution

$$x_k^o = \sum_{j=1}^{k} \alpha_{kj}^o w_j \quad ; \qquad k = 1,\ldots,n$$

with the coefficients α_{nj}^o for which

$$\sum_{k=1}^{n} \lambda_k \sum_{j=1}^{k} (\alpha_{kj}^o - \alpha_{kj}^*)^2 = \sum_{j=1}^{n} \sum_{k=j}^{n} \lambda_k \, (\alpha_{nj}^o - \alpha_{nj}^*)^2 = \min \quad .$$

Obviously, the optimal coefficients α_{nj}^o, $k = j,\ldots,n$ for any $j = 1,\ldots,n$ can be determined from the condition

$$\sum_{k=1}^{n} \lambda_k \, (\alpha_{kj}^o - \alpha_{kj}^*)^2 = \min \quad . \tag{17}$$

Let us fix $j = 1,\ldots,n$. Under the substitution

$$\alpha_{kj} \to \alpha_{k-j+1} \quad ,$$

$$\lambda_k \to \lambda_{k-j+1} \quad , \qquad k = j,\ldots,n,$$

let us consider $\alpha = (\alpha_1,\ldots,\alpha_m)$ as vectors in m-dimensional space with the inner product

$$(\alpha,\beta) : \sum_{k=1}^{m} \lambda_k \alpha_k \beta_n$$

and the metric

$$||\alpha - \beta|| = \sqrt{\sum_{k=1}^{m} \lambda_k \ (\alpha_n - \beta_n)^2} \quad .$$

Let S_m be a simplex of all vectors $\alpha = (\alpha_1,\ldots,\alpha_m)$ which satisfy the constraints (1), namely,

$$\alpha_k \geq 0 \quad , \qquad \sum_{k=1}^{m} \alpha_k \leq 1 \quad . \tag{16'}$$

The problem is to find a vector $\alpha^{\circ} = (\alpha_1^{\circ},\ldots,\alpha_m^{\circ}) \ \varepsilon S_m$ such that

$$||\alpha^{\circ} - \alpha^*|| = \min_{\alpha\varepsilon S_m} ||\alpha - \alpha^*|| \quad . \tag{17'}$$

Of course, $\alpha^{\circ}_{\cdot} = \alpha^*$ if $\alpha^* \ \varepsilon S_m$. Suppose $\alpha^* \notin S_m$. Let us consider a half-line from the point α^*, which is perpendicular to the hyperplane $L = \{\alpha ! \ \sum_{k=1}^{m} \alpha_n = 1 \ _s\}$. It is all points β with coordinates

$$\beta_n = \alpha_k^* - \frac{t}{\lambda_k} \quad ; \qquad k = 1,\ldots,m \ (t \geq 0). \tag{18}$$

Let $\Pi = (\Pi_1,\dots,\Pi_m)$ be a projection of the point α^* onto the hyperplane L:

$$\Pi_k = \alpha_k^* - \frac{t_\Pi}{\lambda_k} , \qquad k = 1,\dots,n$$

$$t_\Pi = \frac{\sum_{k=1}^{m} \alpha_k^* - 1}{\sum_{k=1}^{m} 1/\lambda_k} , \tag{19}$$

and $S'_m = L \cap S_m$ be a set of all vectors $\alpha = (\alpha_1,\dots,\alpha_m)$ such that

$$\sum_{k=1}^{m} \alpha_n = 1 , \qquad \alpha_n \geq 0 .$$

Obviously, if $\Pi \varepsilon S'_m$, i.e.

$$\alpha_j^* \geq \frac{\sum_{k=1}^{m} \alpha_k^* - 1}{\lambda_j \sum_{k=1}^{m} 1/\lambda_k} \qquad \text{for all } j = 1,\dots,m ,$$

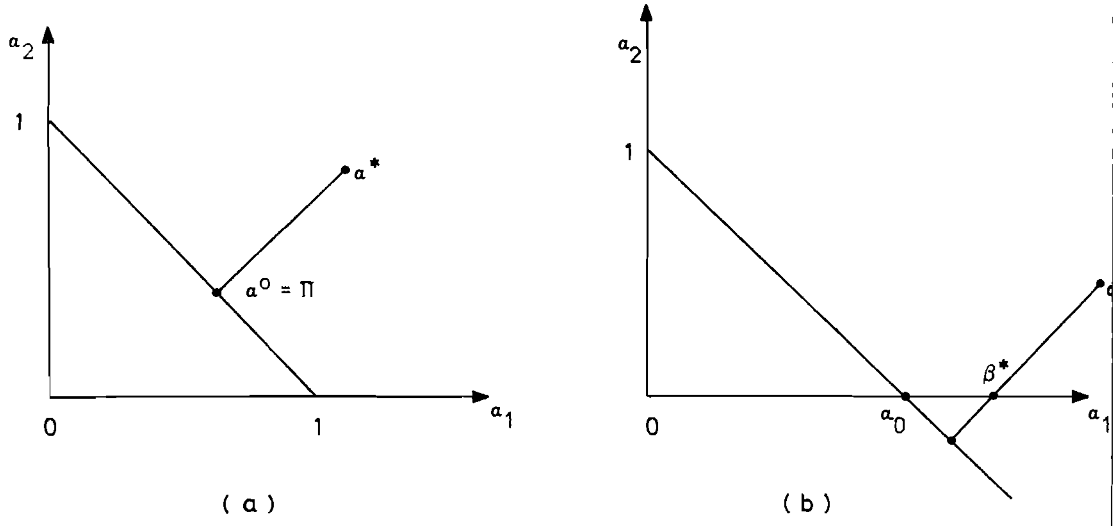
then $\alpha^0 = \Pi$. (See Figure 5.)

Suppose $\Pi \notin S'_m$. This means that some of the coordinates $\Pi_1,\dots,\Pi_m$ are <u>negative</u>. Let us consider

$$\beta_n^* = \alpha_k^* - \frac{t^*}{\lambda_k} , \qquad k = 1,\dots,m , \tag{20}$$

where

$$t^* = \min \lambda_k \alpha_k^* .$$



α_2
1
α^*
$\alpha^0 = \Pi$
0
1
α_1
(a)
α_2
1
β^*
0
α_0
α_1
(b)

FIGURE 5

Obviously, $\sum_{k=1}^{m} \beta_n^* > 1$ and the hyperplane L separates β^* and S_m. Because α^o is the minimum point of

$$\begin{aligned}\min_{\alpha\varepsilon S_m} ||\alpha - \beta||^2 &= \min_{\alpha\varepsilon S_m'} ||\alpha - \beta||^2 \\ &= \min_{\alpha\varepsilon S_m'} ||\alpha - \Pi||^2 + ||\beta - \Pi||^2 \\ &= \min_{\alpha\varepsilon S_m} ||\alpha - \beta||^2 + ||\beta - \Pi||^2\end{aligned}$$

for <u>any</u> β of the type (3), $\sum_{k=1}^{m} \beta_n \geq 1$, we can determine α^o from the condition

$$||\alpha^o - \beta^*|| = \min_{\alpha\varepsilon S_m} ||\alpha - \beta^*|| \quad .$$

From the very beginning we could set coordinates in such a way that

$$\lambda_1\alpha_1^* \geq \cdots \geq \lambda_m\alpha_m^* \quad .$$

Then

$$\min_{1\leq k\leq m} \lambda_k\alpha_k^* = \lambda_m\alpha_m^* \quad ,$$

i.e. $\beta^* = (\beta_1^*,\ldots,\beta_{m-1}^*, 0)$. Obviously,

$$\min_{\alpha\varepsilon S_m} \sum_{k=1}^{m} \lambda_k (\alpha_k - \beta_k^*)^2 = \min_{\alpha\varepsilon S_m} \sum_{k=1}^{m-1} \lambda_k (\alpha_k - \beta_k^*)^2 + \lambda_k\alpha_k$$

$$= \min_{\alpha\varepsilon S_{m-1}} \sum_{k=1}^{m-1} \lambda_k (\alpha_k - \beta_k^*)^2 \quad ,$$

where S_{m-1} denotes the set of all vectors $\alpha = (\alpha_1,\ldots,\alpha_{m-1})$ in (m-1)-dimensional space for which

$$\alpha_k \geq 0 \quad ; \qquad \sum_{k=1}^{m-1} \alpha_n \leq 1 \quad . \tag{16"}$$

Thus $\alpha_m^o = 0$ and the question on $\alpha_1^o,\ldots,\alpha_{m-1}^o$ can be considered in the same way as it was done above, because the problem now is to find a vector $\alpha^o = (\alpha_1^o,\ldots,\alpha_{m-1}^o) \varepsilon S_{m-1}$ such that

$$||\alpha^o - \beta^*|| = \min_{\alpha\varepsilon S_{m-1}} ||\alpha - \beta|| \quad . \tag{17"}$$

Similarly, as it was above, we can determine $\alpha^o = (\alpha_1^o,\ldots,\alpha_{m-1}^o)$ at this <u>second</u> step or reduce our problem to the case of (m-2) unknown components $\alpha_1^o,\ldots,\alpha_{m-2}^o$. Not more than in m steps can we determine all components $\alpha_1^o,\ldots,\alpha_m^o$.

According to the formulas (19) and (20), at every step we have to reduce α_k^*, β_k^* etc. to zero or with subtraction of a value which is proportional to the corresponding $1/\lambda_k$. Thus the optimal $\alpha_1^o,\ldots,\alpha_n^o$ are the following:

$$\alpha_k^o = \max\left(0,\ \alpha_k^* - \frac{t^o}{\lambda_k}\right) \quad ; \qquad k = 1,\ldots,m, \qquad (21)$$

where the constant t^o can be determined from the condition

$$\sum_{k=1}^{m} \alpha_k^o = 1 \quad ; \qquad (22)$$

i.e. there is the crucial level t^o such that

$$\alpha_k^o = \begin{cases} 0 \ , & \text{if } \alpha_k^* \leq \dfrac{t^o}{\lambda_k} \\ \alpha_k^* - \dfrac{t^o}{\lambda_k} \ , & \text{if } \alpha_k^* \geq \dfrac{t^o}{\lambda_k} \ . \end{cases}$$

The same result holds true in the case when, in addition to the constraints (16), we assume--according to (2)--that

$$\alpha_k \leq \alpha_k^* \quad ; \qquad k = 1,\ldots,m,$$

because α_k^o; $k = 1,\ldots,m$ of the type (21) satisfies these "extra" constraints.