# VARIABLE SEPARATION PRINCIPLE FOR MATHEMATICAL PROGRAMMING

IWARO TAKAHASHI

*Waseda University, Tokyo*

(*Received Feb. 4, 1963*)

## § 1. THE CONCEPT OF VARIABLE SEPARATION PRINCIPLE

The ploblem of mathematical programming is to maximize (or minimize) a objective function $f(x)$ subject to constraints, which are usually expressed as $g_i(x)=0$ (or $g_i(x)\geqq 0$), $i=1, 2, \ldots$

We classify these constraints into two types; *simple* and *complicating*. *Simple* constraints are such as $x_i\geqq 0$, $M_i\geqq x_i\geqq 0$, etc., while let us call such constraints *complicating*, that complicate the problem because of involving many variables. In other words, if there were no complicating constraints, variables would be separated into several groups and we could treat the problem separately for each group.

Lagrange function is the sum of objective function and constraint functions weighted with Lagrange multipliers. Kuhn & Tucker (XII) showed the equivalence of mathematical programming and minimax problem of the Lagrange function.

The basic ideas of Variable Separation Principle are as follows;

(a) Select complicating constraints out of given constraints, and take only complicating constraints into the Lagrange function.

(b) Optimize the Lagrange function in $x$ (constrained with simple constraints) with Lagrange multipliers fixed. In this process we can treat variables separately, with appropriately selected complicating constraints.

(c) Next modify the Lagrange multipliers so that complicating constraints are satisfied.

In step (b) the optimal value of the objective function is a function of Lagrange multipliers $\lambda=(\lambda_1, \ldots, \lambda_m)$ which we denote $L(\lambda)$.

In this paper, we clarify the properties of $L(\lambda)$ and make general algorithms for process (c).

The effectiveness of our principle in application depends on the selection to select complicating constraints. We show examples of application of our principle in §5, 6, 7, and 8.

## §2. BELLMAN SURFACE $B(y)$ AND LEGENDRE FUNCTION $L(\lambda)$

Let us consider $x=(x_1, \dots, x_m)$ and its objective function $f(x)$ and constraint functions $g_i(x)$ $(i=1\sim m)$. Besides $g_i(x)=0$ $(i=1\sim m)$ there may be constraints for $x$, the domain restricted with which we denote $C$. $g_i(x)=0$ $(i=1\sim m)$ are correspond to complicating constraints and $C$ is correspond to simple ones.

The maximum value of $f(x)$ subject to $g_i(x)=y_i$ $(i=1\sim m)$ and $x\in C$ is a function of $y=(y_1, \dots, y_m)$, so we denote it $B=B(y)$. That is,

$$(1)\quad B=B(y)=\underset{x\in C}{\mathrm{Max}}\ \{f(x)|\,g_i(x)=y_i\ (i=1\sim m)\}$$

we call it *Bellman surface*, (being suggested by (III) (IV)).

The maximum value of Lagrange function $f(x)+\lambda_1 g_1(x)+\dots+\lambda_m g_m(x)$ in $x\in C$ with $\lambda=(\lambda_1, \dots, \lambda_m)$ fixed, is a function of $\lambda$, so we denote it $L=L(\lambda)$. That is

$$(2)\quad L=L(\lambda)=\underset{x\in C}{\mathrm{Max}}\ \{f(x)+\lambda_1 g_1(x)+\dots+\lambda_m g_m(x)\}$$

We call it *Legendre function*.

(Theorem 1)

If

$$(3)\quad \begin{cases} f(x) \text{ is a convex* function, } g_i(x) \text{ are linear functions } (i=1\sim m) \\ \text{and } C \text{ is a convex set,} \end{cases}$$

then $B(y)$ is a convex function.

Furthermore if $f(x)$ is strictly convex,* then $B(y)$ is strictly convex.

* If $F(\theta x+(1-\theta)y)\geqq\theta F(x)+(1-\theta)F(y)$ $(0\leqq\theta\leqq1)$, we call $F(x)$ convex function, and if $\geqq$ is replaced by $>$, strictly convex.

Let us assume that the conditions (3) are satisfied throughout this paper.

Let $\xi_1, \dots, \xi_m, \phi$ be Cartesian coordinates of $(m+1)$ dimensional space.

(4) $\phi = F \quad (\xi_1, \dots, \xi_m)$

expresses a surface $K$ in this space. (*Point expression of surface*). Consider a supporting plane of $K$ whose normal cosines are proportional to

$$(\eta_1, \dots, \eta_m, 1)$$

which we denote $\pi(\eta_1, \dots, \eta_m)$. The value $\Psi$ of $\phi$-coordinate of the cross point of $\phi$-axis and $\pi(\eta_1, \dots, \eta_m)$ is the function of $\eta = (\eta_1, \dots, \eta_m)$ so denote it

(5) $\Psi = G \quad (\eta_1, \dots, \eta_m)$

Conversely, given (5), a plane whose normal cosines are proportional to $(\eta_1, \dots, \eta_m, 1)$ and which crosses $\phi$-axis at $\Psi = G\ (\eta_1, \dots, \eta_m)$ is uniquely determined in $\phi$-$\xi$ space. The envelope of such planes is the surface $K$. That is, (5) is another expression of $K$ (supporting plane expression of surface), and $G$ is called *Legendre Transform* of $F$ ((V), (VII)). (Fig. 1).

If $F$ is convex, $\pi(\eta_1, \dots, \eta_m)$ is uniquely determined for an $\eta = (\eta_1, \dots, \eta_m)$, and $\xi$-coordinates of contact points are called points (of $\xi$) Legendre-mapped (or simply $L$-mapped) from an $\eta$. If $F$ is strictly convex, there is only one point $L$-mapped from an $\eta$. But generally more than one points may be $L$-mapped from an $\eta$, and a set of these points is a convex set. (Fig. 1). One point of $\xi$ may be simultaneously $L$-mapped from $\eta^1$ and $\eta^2$ $(\eta^1 \neq \eta^2)$. $\pi(\eta_1, \dots, \eta_m)$ is a supporting plane of $K$ at points $L$-mapped from $\eta$.

In the case that if $\xi^i$ is a point $L$-mapped from $\eta^i (i=1, 2)$ and $\eta^1 \neq \eta^2$ then $\xi^1$ allways different from $\xi^2$, a supporting plane of $K$ at a $\xi$ is uniquely determined and coincides with a tangent plane.

Furthermore if (5) is the point expression of $K'$ in $\eta$-$\Psi$ space, its supporting plane expression is (4) (with appropriate direction of the normal).

So (4) is the Legendre Transform of (5).

(Theorem 2)

$L(\lambda)$ is Legendre Transform of $B(y)$, so $L(\lambda)$ is concave.

Let $y$ be any point $L$-mapped from $\lambda$, then

(6) $L(\lambda)=B(y)+\lambda y$

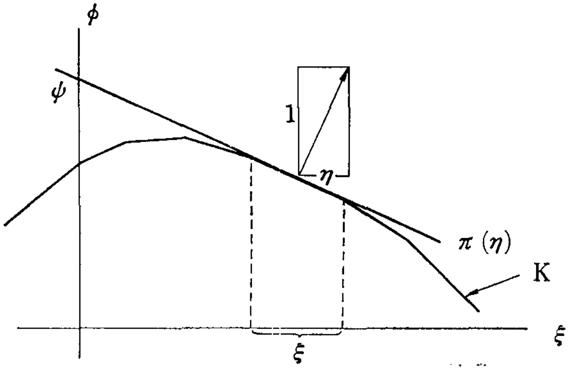



Fig. 1. Legendre Transform

## §3. USE OF $L(\lambda)$ FOR MATHEMATICAL PROGRAMMING

Let us consider a mathematical programming with equality constraints for $g_i(x)$;

(1) Max $f(x)$ Sub $g_i(x)=0$ $(i=1\sim m)$, $x\in C$

*Strictly convex case*

First consider the case where $f(x)$ and $B(y)$ are strictly convex. In this case $x$ maximizing §2 (2) is uniquely determined, so $y$ $L$-mapped from a $\lambda$ is uniquely determined, and the next theorem holds.

(Theorem 1)

Let $\lambda^*$ be a minimum point of $L(\lambda)$, then $y$ $L$-mapped from $\lambda^*$ is

$0=(0,\dots,0)$

Conversely $\lambda$ $L$-mapped from a $y=0$ is $\lambda^*$, and

$$B(0)=L(\lambda^*)$$

Further, $x$ maximizing §2 (2) for $\lambda=\lambda^*$ is the optimal solution of (1).

*General Case*

Next consider the case where $f(x)$ and $B(y)$ are not necessarily strictly conxex.

(Theorem 2)

Let $\lambda^*$ be a minimum point of $L(\lambda)$, then the set of points $L$-mapped from $\lambda^*$ contains $0=(0,\dots,0)$, and

$$B(0)=L(\lambda^*)$$

Furthermore $x$ which maximizes §2 (2) for $\lambda=\lambda^*$ and satisfies $g_i(x)=0$ is a optimal solution of (1).

(Corollary)

If (1) is nonfeasible, then the minimum value of $L(\lambda)$ is not finite.

*Numerical example of a mathematical programming, $B(y)$ and $L(\lambda)$*

Consider a mathematical programming;

$$\text{Max } 3x_1+4x_2 \text{ sub} \begin{cases} x_1+x_2\leqq 3 \\ 2x_1+x_2=2 \\ x_i\geqq 0 \quad (i=1,\ 2) \end{cases}$$

Let

$$f(x)=3x_1+4x_2, \qquad C=\begin{cases} x_1+x_2\leqq 3 \\ x_i\geqq 0 \quad (i=1,\ 2) \end{cases}$$
$$g(x)=2-2x_1-x_2,$$
$$B(y)=\underset{x\in C}{\text{Max}}\ \{3x_1+4x_2 \mid 2-2x_1-x_2=y\}$$

is shown in Fig. 2 (a). (Values of $B$ do not exist for $y>2$, $y<-4$).

$$L(\lambda)=\underset{x\in C}{\text{Max}}\ \{3x_1+4x_2+\lambda\ (2-2x_1-x_2)\}$$

is shown in Fig 2 (b).

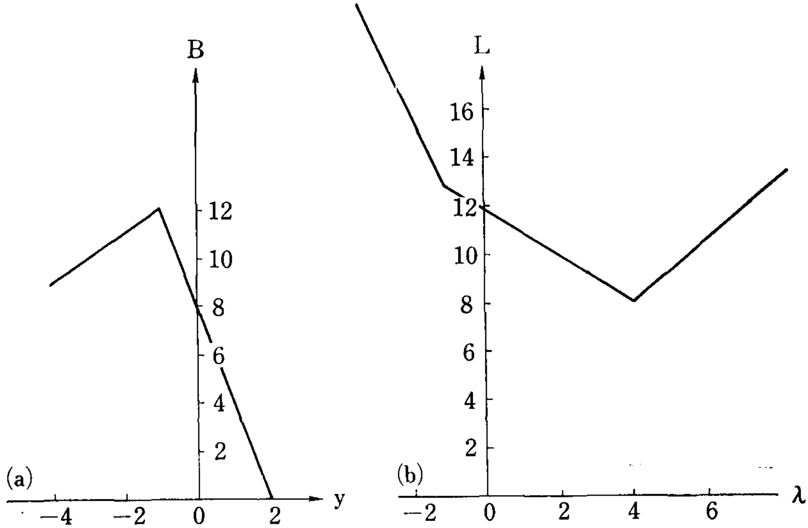



Fig. 2. $B(y)$ and $L(\lambda)$.

## §4. MINIMUM SEARCH OF $L(\lambda)$

*Strictly convex case (local approach)*

Here we explain the successive approximation method to search for a minimum point of $L(\lambda)$ in the case where $f(x)$ and $B(y)$ are strictly convex. For this the next theorem is important, which is an immediate consequence of the properties of Legendre Transform.

(Theorem 1)

If $B(y)$ is strictly convex, $y$ $L$-mapped from a $\lambda$ is uniquely determined, so derivatives of $L(\lambda)$ exist and

(1) $\dfrac{\partial L(\lambda)}{\partial x_i}=y_i \quad (i=1\sim m)$

This theorem means that even if we do not know the explicit form of $L(\lambda)$, we can know the differential coefficients of $L(\lambda)$ for a given $\lambda$.

This leads us to the following algorithm;

*Initial Step*

Chose any $\lambda^0$ for a initial point.

*Iterative Step*

( i ) Let $x$ maximizing §2 (2) for $\lambda=\lambda^\nu$ be $x^\nu$.

(ii) Compute

(2) $y_i^\nu=g_i(x^\nu) \quad (i=1\sim m)$,

and if $y_i^\nu=0$ $(i=1\sim m)$ then $x^\nu$ is the optimal solution of §3 (1) (§3 Theorem 1). Otherwise go to (iii).

(iii) Compute

(3) $\lambda_i^{\nu+1}=\lambda_i^\nu-hy_i^\nu \quad (i=1\sim m),\ h>0,$

and go back to (i).

From (1), (3) we know above algorithm is Gradient method (I) itself. So for sufficiently small $h$ the convergence has been proved with concavity of $L(\lambda)$. We call this method *local' approch.*

*General case* (*Global Approach*)

In the case where $B(y)$ is not necessarily strictly convex, the next theorem corresponds to theorem 1.

(Theorem 2)

$y=(y_1, \dots, y_m)$ $L$-mapped from a $\lambda$ is $y$-coordinate of

(4) $(y_1, \dots, y_m, -1)$

which are propotional to normal cosines of a supporting plane of $L(\lambda)$ at the $\lambda$.

Generally a supporting plane of $L(\lambda)$ at a $\lambda$ is not uniquely determined, so (3) cannot be used for this case. So we propose another method for minimum search. The basic ideas of this method, which are similar to the cutting plane method (XIII), are as follows;

Chose $m+1$ points $\lambda^0, \dots, \lambda^m$ such that the roof* of $\pi^\nu$ $(\nu=0\sim m)$ has

---

* The roof of $n$ planes $L=a_i+a_{i1}\lambda_1+\dots+a_{im}\lambda_m$ $(i=1\sim n)$ is the function (or surface) $\underset{i}{\mathrm{Max}}(a_i+a_{i1}\lambda_1+\dots+a_{im}\lambda_m)$.

a finite minimum value, where $\pi^\nu$ is a supporting plane of $L(\lambda)$ at $\lambda$. Next find the minimum point $\lambda^{m+1}$ of the roof. Let $\pi^{m+2}$ be a supporting plane of $L(\lambda)$ at $\lambda^{m+1}$, and find the minimum point $\lambda^{m+2}$ of the roof of $\pi^\nu$ ($\nu=0\sim m+1$), and so on (Fig. 3). Then as $N\to\infty$, $\lambda^N$ approches a minimum point of $L(\lambda)$. Further, if $L(\lambda)$ is piece-wise linear (like the numerical example of §3), a finite number of iterations suffice.

We call this method *global approach.*

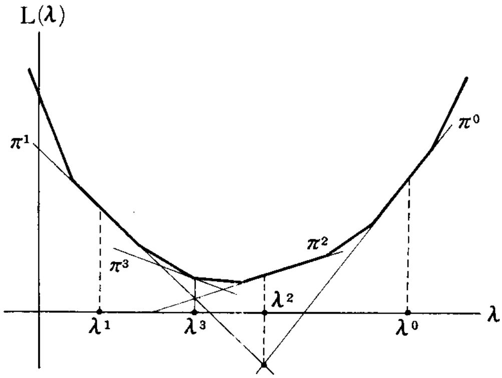



Fig. 3. Global Approch.

Now we describe a theorem by which we concrete above ideas into an algorithm.

(Theorem 3)

Let $\pi^\nu$ be a supporting plane of $L(\lambda)$ at $\lambda^\nu$, and normal cosines of $\pi^\nu$ be proportional to $(y_1^\nu, \dots, y_m^\nu, -1)$ ($\nu=0\sim m$). If $\xi^\nu \geqq 0$ satisfying

$$(5)\quad \xi^0+\xi^2+\cdots+\xi^m=1$$
$$\xi^0 y_i^0+\xi^1 y_i^1+\cdots+\xi^m y_i^m=0 \quad (i=1\sim m)$$

exist, then the roof* of $\pi^\nu(\nu=0\sim m)$ has finite minimum value.

We combine above theorem and the ideas of revised simplex method into the following algorithm;

* See foot note of last page.

*Initial Step*

(i) Choose $\lambda^\nu$ $(\nu=0\sim m)$ so that $y^\nu$ $L$-mapped from $\lambda^\nu$ $(\nu=0\sim m)$. satisfy (5). Compute $B^\nu$, $L^\nu$ for $\lambda^\nu$ $(\nu=0\sim m)$.

(ii) Invert the coefficient matrix of (5).

$$(6)\quad \begin{bmatrix} 1 & 1 & \cdots & 1 \\ y_0^1 & y_1^1 & \cdots & y_1^m \\ \vdots & \vdots & & \vdots \\ y_m^0 & y_m^1 & \cdots & y_m^m \end{bmatrix} = \begin{bmatrix} \xi^0 & \xi^{01} & \cdots & \xi^{m0} \\ \xi^1 & \xi^{11} & \cdots & \xi^{1m} \\ \vdots & \vdots & & \vdots \\ \xi^m & \xi^{m1} & \cdots & \xi^{mm} \end{bmatrix}$$

*Iterative Step*

(i) Compute

$$(7)\quad \begin{cases} \tilde{L}^{m+1}=B^0\xi^0+B^1\xi^1+\cdots+B^m\xi^m \\ \lambda_j^{m+1}=-(B^0\xi^{0j}+B^1\xi^1{}_j+\cdots+B^m\xi^{mj}) \quad (j=1\sim m) \end{cases}$$

$\lambda^{m+1}=(\lambda_1^{m+1},\ldots,\lambda_m^{m+1})$ is a new point. $\tilde{L}^{m+1}$ is a minimum value of the roof of $\pi^\nu$ $(\nu=0\sim m)$.

(ii) Find $y^{m+1}$ $L$-mapped from $\lambda^{m+1}$, and compute $B^{m+1}$, $L^{m+1}$.

$$(8)\quad L^{m+1}=\tilde{L}^{m+1}$$

Then $\lambda^{m+1}$ is a minimum point. Otherwise go to (iii).

(iii) Compute

$$(9)\quad \xi^{i,m+1}=\xi^i+y_1^{m+1}\xi^{i1}+\cdots+y_m^{m+1}\xi^{im} \quad (i=1\sim m),$$

and take $\xi^{i,m+1}$ into $i$-th element of $(m+1)$-th column of (6), then resulting matrix is $(m+1)\times(m+2)$. Let $i_0$ be a number $i$ which attains

$$(10)\quad \min_i \xi^i/\xi^{i,m+1} \quad (\xi^{i,m+1}>0)$$

Sweep out the above $(m+1)\times(m+2)$ matrix with pivot$\xi^{i_0,m+1}$, then delete the $(m+1)$-th column. Take number $m+1$ for $i_0$. Then return to (i).

Through above stop following relation, holds;

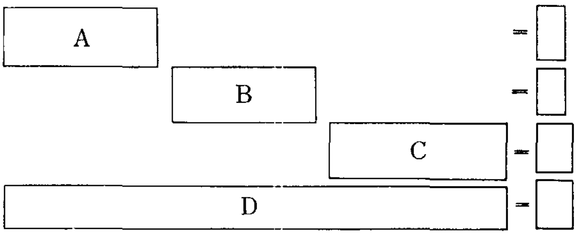
$$(11)\quad \text{Min}\,(L^0,\ldots,L^m)\geqq \text{Min}\ L(\lambda)\geqq \tilde{L}^{m+1}.$$

(11) is usefull to determine the range of Min $L(\lambda)$ when we stop the computation half-way.

General algorithm to find initial points was not stated, which we must device for individual examples.

## §5. LINEAR PROGRAM OF BLOOK DIAGONAL TYPE

Consider linear programs with constraints shown in the following diagram. (VI)




If we select the constraints of D part as complicating, then variables are separated into A, B, C part. In this case $f(x)$ is not strictly convex so we use global approch.

We trace the process along the next example;
Two groups of variables $x_{ij}^{(1)}, x_{ij}^{(2)}$ are respectively subject to the following transportation type equality constraints and $x_{ij}^{(1)} \geqq 0, x_{ij}^{(2)} \geqq 0$;

(1)

| | | | |
|---|---|---|---|
| $x_{11}^{(1)}$ | $x_{12}^{(1)}$ | $x_{13}^{(1)}$ | 12 |
| $x_{21}^{(1)}$ | $x_{22}^{(1)}$ | $x_{23}^{(1)}$ | 15 |
| 7 | 10 | 10 | 27 |

| | | | |
|---|---|---|---|
| $x_{11}^{(2)}$ | $x_{12}^{(2)}$ | $x_{13}^{(2)}$ | 35 |
| $x_{21}^{(2)}$ | $x_{22}^{(2)}$ | $x_{23}^{(2)}$ | 40 |
| 20 | 25 | 30 | 75 |

Cost coefficient $c_{ij}^{(k)}$ are

(2)

$c_{ij}^{(1)}$

| | | |
|---|---|---|
| 1 | 2 | 3 |
| 2 | 4 | 6 |

$c_{ij}^{(2)}$

| | | |
|---|---|---|
| 2 | 3 | 4 |
| 3 | 6 | 8 |

Further $x_{ij}^{(1)}$ $x_{ij}^{(2)}$ are subject to another constraint

(3) $g(x)=720-\sum_k\sum_i\sum_j a_{ij}^{(k)}x_{ij}^{(k)}=0,$

where

(4)

$a_{ij}^{(1)}$

| | | |
|---|---|---|
| 3 | 5 | 7 |
| 4 | 6 | 7 |

$a_{ij}^{(2)}$

| | | |
|---|---|---|
| 6 | 7 | 8 |
| 8 | 8 | 8 |

We want to maximize

(5) $f(x)=-\sum_k\sum_j\sum_i c_{ij}^{(k)}\ x_{ij}^{(k)}.$

The coefficients $C'$ of

(6) $-\{f(x)+\lambda g(x)\}$

can be written as

(7)

| | | |
|---|---|---|
| $1+3\lambda$ | $2+5\lambda$ | $3+7\lambda$ |
| $2+4\lambda$ | $4+6\lambda$ | $6+7\lambda$ |

| | | |
|---|---|---|
| $2+6\lambda$ | $3+7\lambda$ | $4+8\lambda$ |
| $3+8\lambda$ | $6+8\lambda$ | $8+8\lambda$ |

Under the condition with $\lambda$ fixed, we have only to solve separately two transportation problems with cost coefficients (7) and constraints (1).

*Initial Step*

(i) Let $\lambda^0=0$ (arbitrary value of $\lambda$), and find $x^0, y^0, B^0, L^0$ (Table 0). $y^0$ is negative so $\lambda^1$ must be a positive lage value, but from (7) we know $\lambda^1=3$ suffice. For $\lambda^1=3$ find $x^1, y^1, B^1, L^1$ (Table 1).

(ii) Invert

$$\begin{bmatrix} 1 & 1 \\ y^0 & y^1 \end{bmatrix}=\begin{bmatrix} 1 & 1 \\ -31 & 29 \end{bmatrix}$$

into

(8) $$\begin{bmatrix} \xi^0 & \xi^{01} \\ \xi^1 & \xi^{11} \end{bmatrix}=\begin{bmatrix} 29/60 & -1/60 \\ 31/60 & 1/60 \end{bmatrix}$$

We follow the general algorithm of §4.

*Iterative Step*

(i)

| | | | | |
|---|---|---|---|---|
| −395 | ⓪ | 29/60 | −1/60 | $\tilde{L}^2=-437.2$ |
| −477 | ① | 31/60 | 1/60 | $\lambda^2=\ 41/30$ |
| | | −437.2 | −41/30 | |

(ii) Find $x^2, y^2, B^2, L^2$ for $\lambda^2=41/30$ (Table 2).

(iii)

| | 1 | −3 | | |
|---|---|---|---|---|
| ⓪ | 29/60 | −1/60 | (32/60) | |
| ① | 31/60 | 1/60 | 29/60 | ↓ Sweep Out (with pivot ( )) |
| | 29/32 | −1/32 | | |
| | 3/32 | 1/31 | | |

(i′)

| | | | | |
|---|---|---|---|---|
| −423 | ⓪ | 29/32 | −1/32 | $\tilde{L}^3=-428.1$ |
| −477 | ① | 3/32 | 1/32 | $\lambda^3=27/16$ |
| | | −428.1 | −27/16 | |

(ii′) Find $x^3, y^3, B^3, L^3$ for $\lambda^3=27/16$ (Table 3).

(iii′)

| | 1 | 17 | | |
|---|---|---|---|---|
| ② | 29/32 | −1/32 | 12/32 | |
| ① | 3/32 | 1/32 | (20/32) | ↓ Sweep Out (with pivot ( )) |
| ② | 17/20 | −1/20 | | |
| ③ | 3/20 | 1/20 | | |

(i″)

| | | | | |
|---|---|---|---|---|
| −423 | ② | 17/20 | −1/20 | $\tilde{L}^4=-427.5$ |
| −453 | ③ | 3/20 | 1/20 | $\lambda^4=3/2$ |
| | | −427.5 | −3/2 | |

(ii″) Find $x^4, y^4, B^4, L^4$ for $\lambda^4=3/2$ (Table 4).

$$L^4=\tilde{L}^4$$

so $\lambda^4=3/2$ is a minimum point.

*Final Treatment*

If in the final Table $x_{ij}^{(k)}$ are unique, then this is the optimal solution (§3 Theorem 2). Otherwise, we can choose optimal $x_{ij}^{(k)}$ out of undetermined $x_{ij}^{(k)}$ in final Table (§3 Theorem 2). That is; $x_{ij}^{(2)}$ circled in in Table 4 are determined by $6x_{11}^{(1)}+8x_{21}^{(2)}+8x_{13}^{(2)}+8x_{23}^{(2)}=397$ and constraints (1). The results are

$x_{ij}^{(1)}$

| | | |
|---|---|---|
| 0 | 10 | 2 |
| 7 | 0 | 8 |

$x_{ij}^{(2)}$

| | | |
|---|---|---|
| 1.5 | 25 | 8.5 |
| 18.5 | 0 | 21.5 |

Table 0

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $C'$ | 1 | 2 | 3 | 2 | 3 | 4 | |
| | 2 | 4 | 6 | 3 | 6 | 8 | |
| $x$ | | 2 | 10 | | 5 | 30 | $B^0=-395,\ L^0=-395$ |
| | 7 | 8 | | 20 | 20 | | $y^0=-31$ |

Table 1

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $C'$ | 10 | 17 | 24 | 20 | 24 | 28 | $\lambda^1=3$ |
| | 14 | 22 | 27 | 27 | 30 | 32 | |
| $x$ | 2 | 10 | | 20 | 15 | | $B^1=-477,\ L^1=390$ |
| | 5 | | 10 | | 10 | 30 | $y^1=29$ |

Table 2

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $C'$ | 153 | 265 | 377 | $\frac{1}{30}$ | 306 | 377 | 448 | $\frac{1}{30}$ $\lambda^2=41/30$ |
| | 224 | 366 | 467 | | 418 | 508 | 568 | |
| $x$ | | 10 | 2 | | | 25 | 10 | $B^2=-423,\ L^2=-427.1$ |
| | 7 | | 8 | | 20 | | 20 | $y^2=-3$ |

Table 3

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $C'$ | 97 | 167 | 237 | $\frac{1}{16}$ | 194 | 237 | 280 | $\frac{1}{16}$ | $\lambda^3=27/16$ |
| | 140 | 226 | 285 | | 264 | 312 | 344 | | |
| $x$ | | 10 | 2 | | 10 | 25 | | | $B^3=-453,\ L^3=-428.1$ |
| | 7 | | 8 | | | | 30 | | $y^3=17$ |

Table 4

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $C'$ | 11 | 19 | 27 | $\frac{1}{2}$ | 22 | 27 | 32 | $\frac{1}{2}$ | $\lambda^4=3/2$ |
| | 16 | 26 | 23 | | 30 | 36 | 40 | | |
| $x$ | | 10 | 2 | | ⑤ | 25 | ⑤ | | $B^4=-438,\ L^4=-427.5$ |
| | 7 | | 8 | | ⑮ | | ㉕ | | $(y^4=7)$ |

## §6. RESOURCE ALLOCATION

Let us call the following problem *Resource Allocation*;

(1) $f(x)=-\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \longrightarrow$ Max.

(2) $\sum_{j=1}^{n} x_{ij}=a_i \quad (i=1\sim m) \quad (a_i\geqq 0)$

(3) $\sum_{i=1}^{m} r_{ij}x_{ij}=b_j \quad (j=1\sim n) \quad (b_j\geqq 0,\ r_{ij}\geqq 0)$

(4) $x_{ij}\geqq 0 \quad (i=1\sim m,\ j=1\sim n)$

We select (2) as complicating. Let

(5) $g_i(x)=a_i-\sum_{j=1}^{n} x_{ij} \quad (i=1\sim m)$

(6) $f(x)+\lambda_1 g_1(x)+\cdots+\lambda_m g_m(x)$

$$=-\sum_{j=1}^{n}\{(c_{1j}+\lambda_1)x_{1j}+\cdots+(c_{mj}+\lambda_m)x_{mj}\}+\lambda_1 a_1+\cdots+\lambda_m a_m$$

In maximizing (6) subject to (3) (4), we can treat variables separately for each $j$. For each $j$, let $I_j$ be a set of number $i$ attaining

(7) $\text{Min}\left\{\frac{\lambda_1+c_{1j}}{r_{1j}},\ldots,\frac{\lambda_m+c_{mj}}{r_{mj}}\right\}$,

then $x_{ij}$ satisfying (8) maximize (6);

$$(8)\quad \begin{cases} \sum_{i\in I_j} x_{ij}=b_j \\ x_{ij}=0 \text{ for } i\notin I_j \end{cases}$$

and maximum value of $L(\lambda)$ is

$$(9)\quad L(\lambda)=-\sum_j b_j \operatorname{Min}\left\{\frac{\lambda_1+c_{1j}}{r_{1j}},\dots,\frac{\lambda_m+c_{mj}}{r_{mj}}\right\}+\sum_i a_i\lambda_i$$

(We need not express $L(\lambda)$ explicitly for our pourpose. But by (9) we know $L(\lambda)$ is a piece-wise linear concave function).

We must use global approach to for search a minimum of $L(\lambda)$. Let us treat the following example:

$c_{ij}$

| | | |
|---|---|---|
| 4 | 6 | 7 |
| 5 | 3 | 17 |

| | $r_{ij}$ | | | $a_i$ |
|---|---|---|---|---|
| | 1 | 2 | 3 | 9 |
| | 2 | 1 | 4 | 15 |
| $b_j$ | 13 | 15 | 30 | |

*Initial Step*

(i) Let $\lambda_1^0=0$, $\lambda_2^0=0$, and maximize (6) for $\lambda=\lambda^0$. For this purpose we have only to write down the Table typed (7), circle the minimum in each column, and fill up the entries corresponding to circles in $x$ Table (Table A0).

$y_1^0$ is negative so we choose a large positive value as $\lambda_1^1$, 6 will suffice. So let $\lambda_1^1=6$, $\lambda_2^1=0$, and maximize (6) for $\lambda=\lambda_1^1$. By the same procedure, we get Table A1. Next choose $\lambda_1$, $\lambda_2$ so that $y_1$, $y_2$ are positive. Let $\lambda_1^2=18$, $\lambda_2^2=12$ and we get Table A2.

(ii) Invert

$$\begin{bmatrix} 1 & 1 & 1 \\ y_1^0 & y_1^1 & y_1^2 \\ y_2^1 & y_2^1 & y_2^2 \end{bmatrix}=\begin{bmatrix} 1 & 1 & 1 \\ -8.5 & 9 & 1.5 \\ 8.5 & -14 & 1.0 \end{bmatrix}$$

into Table B1.

*Iterative Step*

(i) We get $\tilde{L}^3$, $\lambda_1^3$, $\lambda_2^3$ by inner product of $B^0$, $B^1$, $B^2$ and each column of Table B1.

(ii) Maximize (6) for $\lambda_1^3=9.2$, $\lambda_2^3=4.6$ (Table A3). $L^3=\tilde{L}^3$ so $\lambda^3$ is a minimum point.

*Final Treatment*

Fróm Table A3, we solve

$$x_{12}+x_{13}=9 \qquad 2x_{12}+x_{22}=15$$

$$x_{22}+x_{23}=15-6.5 \qquad 3x_{13}+4x_{23}=30$$

and optimal solution is

| | | |
|---|---|---|
| 6.5 | 5.8 | 3.2 |
| 0 | 3.4 | 5.1 |

Table A0

| $\lambda^0$ | | | | | 13 | 15 | 30 | $y^0$ |
|---|---|---|---|---|---|---|---|---|
| 0 | 4/1 | (6/2) | 7/3 | 9 | | (7.5) | 10 | −8.5 |
| 0 | (5/2) | (3/1) | 17/4 | 15 | 6.5 | ( ) | | 8.5 |

$B^0$ : −147.5 $L^0$ : −96.5

Table A1

| $\lambda^1$ | | | | | 13 | 15 | 30 | $y^1$ |
|---|---|---|---|---|---|---|---|---|
| 6 | 10/1 | 12/2 | 13/3 | 9 | | | | 9 |
| 0 | (5/2) | (3/1) | (17/4) | 15 | 6.5 | 15 | 7.5 | −14 |

$B^1$ : −205 $L^1$ : −151

Table A2

| $\lambda^2$ | | | | | 13 | 15 | 30 | $y^2$ |
|---|---|---|---|---|---|---|---|---|
| 18 | 22/1 | (24/2) | 25/3 | 9 | | 7.5 | | 1.5 |
| 12 | (17/2) | 15/1 | (19/4) | 15 | 6.5 | | 7.5 | 1.0 |

$B^2$ : −205 $L^2$ : −166

Table A3

| $\lambda^3$ | | | | | | | | $y^3$ |
|---|---|---|---|---|---|---|---|---|
| 9.2 | 13.2/1 | (15.2/2) | (16.2/3) | 9 | | (7.5) | (10) | −8.5 |
| 4.6 | (9.6/2) | (7.6/1) | (21.6/4) | 15 | 6.5 | ( ) | ( ) | 8.5 |

$B^3$ : −147.5 $L^3$ : −186.6

Table Bl.

| | | | | |
|---|---|---|---|---|
| −147.5 | ⓪ | .3200 | −.1600 | −.0800 |
| −205 | ① | .2266 | −.0800 | −.1066 |
| −205 | ② | .4534 | .2400 | .1866 |

## § 7. RESOURCE ALLOCATION UNDER UNCERTAIN DEMANDS

Here we consider Resource Allocation under uncertain demands which was discussed in (VIII) etc. Here we formulate it as follows; Constraints are

(1) $a_i=\sum_{j=1}^{n} x_{ij}\ (i=1\sim m)$

(2) $y_j=\sum_{i=1}^{m} r_{ij}x_{ij}\ (j=1\sim n)$

(3) $x_{ij}\geqq 0,$

and objective function to be maximized is

(4) $\sum_j F_j(y_j)-\sum_i\sum_j c_{ij}x_{ij}$

where

$$F_j(y)=p_j\left\{\int_0^y sf_j(s)ds+y\int_y^\infty f_j(s)ds\right\}.$$

$f_j(s)$ is a distribution density function of demands. It is easy to know that (4) is convex function of $x_{ij}$ but is not strictly convex.

We take (1) as complicating constraints, and Lagrange function is

(5) $\sum_j F_j(y_j)-\sum_i\sum_j c_{ij}x_{ij}+\sum_i\lambda_i(a_i-\sum_j x_{ij})$

In maximizing (5) under the condition with $\lambda$ fixed, we can treat variables separately for each $j$.

For each $j$ we have only to maximize

(6) $F_j(y_j)-\sum_i(c_{ij}+\lambda_i)x_{ij}$

subject to (2) (3). The contour of

$$F_j(y_j)=F_j(\sum_i r_{ij}x_{ij})$$

is linear, so only one of $x_{ij}$ maximizing (6) is positive (others are all zero.) Therefore, let $\bar{x}_{ij}$ be a root of

(7) $F_j'(r_{ij}x_{ij})r_{ij}-(c_{ij}+\lambda_i)=0,$

then

$$(8)\quad \operatorname*{Max}_{i}\{F_j(r_{ij}\bar{x}_{ij})-(c_{ij}+\lambda_i)x_{ij}\}$$

is the desired one. And let $i_0$ be a number $i$ attaining (8), the $\bar{x}_{i_0j}$ is a solution.

For the minimum search of $L(\lambda)$ we must use a global approach.

## § 8. MULTISTAGE ALLOCATION

Here we consider the Multistage Allocation proposed by R. Bellman (II). We formulate it as follows;

Maximize

$$(1)\quad \sum_{k=1}^{N} f(x_k^1, \ldots, x_k^n)$$

Snbject to

$$(2)\quad \begin{cases} X_0 = x_1^1 + \ldots + x_1^n \\ a_1x_1^1 + \ldots + a_nx_1^n = x_2^1 + \ldots + x_2^n \\ a_1x_2^1 + \ldots + a_nx_2^n = x_3^1 + \ldots + x_3^n \\ \cdot\cdot \\ a_1x_{N-1}^1 + \ldots a_nx_{N-1}^n = x_N^1 + \ldots + x_N^n \\ a_1x_N^1 + \ldots + a_nx_N^n = X_N \end{cases}$$

($X_0$, $X_N$ are given constants)

$$(3)\quad h^i(x_k^1, \ldots, x_k^n) \leqq 0 \quad (i=1\sim m,\ k=1\sim N)$$

(The value of each equation in (2) is called state variable in Dynamic Programming method)

Select (2) as complicating constaints, and construct the Larange function then

$$
(4)\quad \begin{cases} f(x_1^1 \ldots x_1^n) - \lambda_0(x_1^1 + \cdots + x_1^n) + \lambda_1(a_1 x_1^1 + \cdots + a_n x_1^n) \\ + f(x_2^1 \cdots x_2^n) - \lambda_1(x_2^1 + \cdots + x_2^n) + \lambda_2(a_1 x_2^1 + \cdots + a_n x_2^n) \\ \ldots \\ + f(x_N^1 \ldots x_N^n) - \lambda_{N-1}(x_N^1 + \cdots + x_N^n) + \lambda_N(a_1 x_N^1 + \cdots + a_n x_N^n) \\ + \lambda_0 X_0 - \lambda_n X_N \end{cases}
$$

Let $L(\lambda_{k-1}, \lambda_k)$ be the maximum value of

$$
(5)\quad \begin{cases} f(x_k^1 \ldots x_k^n) - \lambda_{k-1}(x_k^1 + \cdots + x_k^n) + \lambda_k(a_1 x_k^1 + \cdots + a_n x_k^n) \\ \text{subject to} \\ h^i(x_k^1 \cdots x_k^n) \leqq \quad (i = 1 \sim m) \end{cases}
$$

then

$$
(6)\quad L(\lambda) = \lambda_0 X_0 + L(\lambda_0, \lambda_1) + \cdots + L(\lambda_{N-1}, \lambda_N) - \lambda_N X_N
$$

In this case we can minimize $L(\lambda)$ by recurrence relations. First we consider the case where $f(x^1 \ldots x^n)$ is strictly convex. Then $L(\xi, \eta)$ has derivatives (§ 4 Theorem 1) and

$$
(7)\quad \begin{cases} X_0 + L_\xi(\lambda_0, \lambda_1) = 0 \\ L_\eta(\lambda_0, \lambda_1) + L_\xi(\lambda_1, \lambda_2) = 0 \\ \ldots \\ L_\eta(\lambda_{N-1}, \lambda_N) - X_N = 0 \end{cases}
$$

are conditions for a minimum point of $L(\lambda)$.
To solve (7);

First fix $\lambda_0$, and then determine $\lambda_1$ by the first equation of (7). Next determine $\lambda_2$ by the second equation of (7), and so on. Let $F(\lambda_0)$ be the value of left hand side of the last equation of (7). Control $\lambda_0$ for $F(\lambda_0)$ to be zero. (for example, by Regula Falsi method)

Above procedures are generalized to case where $f(x^1 \ldots x^n)$ is not necessarily strictly convex. Let

$$Y_0,\ Y_1,\ \dots,\ Y_N$$

be the difference between the left hand side and right hand side of (2).

First fix $\lambda_0$, and then determine $\lambda_1$ so that $Y_0$ is zero in (5) ($k=1$). Next determine $\lambda_2$ so that $Y_1$ is zero in (5) ($k=2$), and so on. Modifying method of $\lambda_0$ is the same as in strictly concave case.

Above method is applicable for high dimensional state variable cases. The method of recurrence functional relation by Bellman (II) reveals weak points in high dimensional case.

The method need vast number of iterations and vast number of memories in high dimensional case. Our method will delete these difficulties.

The reason Pontryagin's Maximum Principle is more powerfull than Bellman's Principle in the problem of automatic control process, is that the former uses the function $\psi(t)$ which is a solution of adjoint equations of equations satisfied by variations of state variables. $\lambda_k$ which we introduced in this §, are corresponds to $\psi(t)$.

## § 9. DISCUSSION

Our principle is similar to Dantzig's Decomposition Principle (VI). Both use Lagrange multipliers to separate variables. Dantzig's method utilize only total value of $y$.

Our method can be applied to Transportation Problem, and in this application will be similar to Primal Dual Algorithm (IX) (X) (X), the effects of which depends on being able to solve Primal Problem very simply through network flow properties. But if non-unity coefficients are involved (such as in resource Allocation), this properties do not hold.

The defect of global approach in this paper is that the general algorithm to find $m+1$ initial points is not shown. To construct this algorithm is left for us. To generalize our methods to inequality constraints for $g_i(x)$ is also left for us.

## § 10. PROO OF THEOREMS

(i) §2 *Theorem 1*

Let $x^1$ be a $x$ attaining

(1) $B(y^1)=\underset{x\in C}{\mathrm{Max}}\{f(x)|\, g(x)=y^1\}$

Let $x^2$ be a $x$ attaining

(2) $B(y^2)=\underset{x\in C}{\mathrm{Max}}\{f(x)|\, g(x)=y^2\}$

For any $\alpha^1$, $\alpha^2$ ($\alpha^1\geqq 0$, $\alpha^2\geqq 0$, $\alpha^1+\alpha^2=1$),

let $\bar{x}$ be a $x$ attaining

(3) $B(\alpha^1y^1+\alpha^2y^2)=\underset{x\in C}{\mathrm{Max}}\{f(x)|\, g(x)=\alpha^1y^1+d^2y^2\}$,

$$g(\alpha^1x^1+\alpha^2x^2)=\alpha^1y^1+\alpha^2y^2 \text{ (for } g \text{ is linear)},$$

and $C$ is convex, so $\alpha^1x^1+\alpha^2x^2$ statisfies the constraints in (3).

Therefore

(4) $f(\bar{x})\geqq f(\alpha^1x^1+\alpha^2x^2)$

and, because of convexity of $f$,

(5) $f(\alpha^1x^1+\alpha^2x^2)\geqq\alpha^1f(x^1)+\alpha^2f(x^2)$.

From (4) (5) (1) (2) (3) we have

(6) $B(\alpha^1y^1+\alpha^2y^2)\geqq\alpha^1B(v^1)+\alpha^2B(y^2)$

So $B(y)$ is convex. Furthermore, if $f(x)$ is stictly convex $\geqq$ in (5) becomes $>$ so $\geqq$ in (6) becomes $>$. (Here $g$ is a vector form of ($g_1$, ..., $g_m$))

(ii) §3 *Theorem 2*

Any point in $y$-$B$ space with

(1) $y=g(x)$, $B=f(x)$, $x\in C$

is on or below the Bellman surface $B=B(y)$.

Let $x^0$ be a $x$ attaining § 2 (2) for a $\lambda$. $y^0$, $B^0$ for $x^0$ by (1) is on the Bellman surface. A plane through $y^0$, $B^0$ with normal cosines proportional to

$$(\lambda_1, \dots, \lambda_m, 1)$$

(2) $B-B^0+\lambda(y-y^0)=0$

is a supporting plane of $B=B(y)$. From §2 (2) we have

(3) $B^0+\lambda y^0=L$

which is the value of $B$ in (2) for $y=0$, so is $B$-coordinate of the cross point of $B$-axis and the supporting plane.

The concavity of $L(\lambda)$ is a immediate consequence of the convexity of $B(y)$. §2 (6) is (3) itself.

(iii) § 4 *Theorem 3*

Let $(\lambda^*, L^*)$ be a point which is simultaneously on planes $\pi^\nu(\nu=0\sim m)$.

The value of the roof function for any direction $r=(r_1, \dots, r_m)$ and a small $\varepsilon>0$ is

(1) $\operatorname*{Max}_\nu\{y^\nu\cdot(\lambda^*+\varepsilon r)+B^\nu\}=L^*+\varepsilon\operatorname*{Max}_\nu(y^\nu\cdot r)$

Now we have

(2) $\sum_{\nu=0}^{m}\xi^\nu y^\nu=0$, so

(3) $\sum_{\nu=0}^{m}\xi^\nu(r\cdot y^\nu)$.

$\xi^\nu\geqq 0$ are all not 0 so at least one of $r\cdot y^\nu$ is $\geqq 0$. Therefore $\operatorname*{Max}_\nu (y^\nu\cdot r)$ is $\geqq 0$, and $\lambda^*$ is a local minimum point of the roof function. The roof function is concave, so its local minimum is also a global minimum.

(vi) *A Global Approach*

Here we prove that $\lambda^{m+1}$ determined by §4 (7) is a minimum point of the roof of $\pi^\nu(\nu=0\sim m)$. The equation $\pi^\nu$ is

(1) $L=B^{\nu}+y_1{}^{\nu}\lambda_1+\cdots+y_m{}^{\nu}\lambda_m \quad (\nu=0\sim m)$,

so $\lambda$ satisfing (1) is determined by

(2) $$\begin{bmatrix} L \\ -\lambda_1 \\ \vdots \\ -\lambda_m \end{bmatrix} = \begin{bmatrix} 1 & y_1{}^0 & \cdots & y_m{}^0 \\ \vdots & \vdots & & \vdots \\ 1 & y_1{}^m & \cdots & y_m{}^m \end{bmatrix}^{-1} \begin{bmatrix} B^0 \\ \vdots \\ B^m \end{bmatrix}$$

It is clear that $\lambda$ satisfying (1) is a minimum point of the roof. q.e.d.

From the fact that the matrix produced by §4 Iterative Step (iii) is §4 (6) type matrix with directional coefficients $y^{\nu}$ of $\pi^{\nu}$ ($\nu\in\{0, 1, \ldots, m\}-i_0+(m+1)$), and §4 (10), $\xi^{\nu}\geqq 0$ is preserved in each iteration.

## REFERENCES


(I) K.J. Arrow & L. Hurwitz: "Gradient Methods for Constrained Maximum" JORSA Vol. 5, No. 2 (1957).

(II) R. Bellman: "Dynamic Programming" Princeton University Press, Princeton, New Jersey, (1957).

(III) R. Bellman & S. Dreyfus: "Applied Dynamic Programming" Princeton University Press, (1962).

(IV) R. Bellman: "Dynamic Programming and Lagrange-multipliers" Proc. Nat. Acad. Sci. U.S.A., Vol. 42, (1956) p. 767–769.

(V) R. Courant & D. Hilbert: "Method en der Mathematischen Physik" Berlin Verlag von Julius Springer, (1931)

(VI) G.B. Dantzig & P. Wolfe: "A decomposition principle for linear programs" The RAND Corporation Paper P-1544, Nov. 1958; revised Nov. 1959, and "G.B. Dantzig Hakushi Koen Kiroku — Generalized Linear Programming" Suugaku (Japan Mathematical Association, Iwanami) Vol. 11, No. 4, (1960) p. 238–240.

(VII) J.B. Dennis: "Mathematical Programming and Electrical Networks" MIT Press and Wiley, (1959).

(VIII) S.E. Elmaghraby: "Allocation under Uncertainly when the Demands has continuous Distribution Function" Management Science Vol. 6, (1960) No. 3, p. 270–294.

(IX) L.R. Ford & D.R. Fulkerson: "A Simple Algorithm for Finding Maximal Network Flows and an Application to the Hitchcock Problem" Canadian Journal of Mathematics Vol. 9, (1957) p. 210–218.

(X) L.R. Ford & D.R. Fulkerson: "A Primal Dual Algorithm for the Capacitated Hitchcock Problem" Naval Research Logistic Quarterly Vol. 4, No. 1, (1957) p. 47–.

(XI) G.B. Dantzig, L.R. Ford & D.R. Fulkerson; "A Primal-Dual Algorithm for Linear Programs" Linear Inequalities and Related Systems, Annuals of Mathematics Study 38, Princeton Univ. Press, (1956), p. 171–181.

(XII) H.W. Kuhn & A.W. Tucker: "Nonliear Programming" Proc. Second Barkely Symposium on Math. Stat. and Prof., Univ. of California Press, Berkeley, (1951) p. 481–482.

(XIII) P. Wolfe: "Some Simplex-like Non-linear Programming Procedures" JORSA Vol. 10, No. 4, (1962) p. 438–447.

(XIV) L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, & E.F. Mishchenko: "The Mathematical Theory of Optimal Process" Wiley, (1962).