Three Algorithms for a Simple Nonlinear Programming Problem

Ilya V. Gouevsky

December 1974 WP-74-73

Three Algorithms for a Simple Nonlinear Programming Problem

Ilya V. Gouevsky

In IIASA working paper WP-74-61[1] Yu. A. Rozanov suggested an elegant method for solving the following optimization problem:

$$(1) \qquad u(x) = \max \sum_{i=1}^{n} u_i(x_i)$$

subject to

$$(2) \qquad \sum_{i=1}^{n} x_i = y, (x_i \geq 0) \quad ,$$

where $u_i(x_i)$ are concave utility functions.

For solving of this problem, when $x_i \geq \tau$, τ is a small number approaching zero, other computational procedures could be suggested. They follow closely the idea of parametrical solution proposed in [1].

Let us divide all indexes of the variables x_i into two subsets: M_1 and M_2, $M_1 \cup M_2 = M = \{1,\ldots,k,\ldots,n\}$; $M_1 \cap M_2 = \emptyset$.

Suppose, the first derivatives

$$(3) \qquad u_i'(x_i) = \lambda \text{ for all } i \in M_1$$

(λ is still an unknown parameter) and the inverse function of (3) exist, i.e.

$$(4) \qquad x_i = \varphi_i(\lambda) \quad , \qquad \text{for all } i \in M_1 \quad .$$

Then a following theorem holds true.

Theorem. If for a given $\lambda = \lambda_1$, the value $u_i'(\tau) \leq \lambda_1$, $i \varepsilon M_2$, then vector $x^* = \{x_i, i\varepsilon M\}$ obtained by the expression

$$(5) \qquad x_i^* = \begin{cases} \varphi_i(\lambda) = \nu_i , & i \varepsilon M_1 \\ \varphi_i(\tau) = \tau , & i \varepsilon M_2 \end{cases}$$

is optimal vector (maximizing (1) subject to (2)).

Proof. For proving the theorem it is sufficient to be shown that

$$(6) \qquad \Delta = u_i(x_i^*) - u_i(x_i) = \sum_{i\varepsilon M_1} u_i(\nu_i) + \sum_{i\varepsilon M_2} u_i(\tau) - \sum_{i\varepsilon M} u_i(x_i) \geq 0 ,$$

for any vector x satisfying (2).

Since the function $\sum_{i\varepsilon M} u_i(x_i)$ is concave, it is twice differentiable and it can be expanded in Taylor's series in the neighborhood of its optimum, i.e.

$$(7) \qquad \sum_{i\varepsilon M} u_i(x_i) = \sum_{i\varepsilon M_1} u_i(\nu_i) + \sum_{i\varepsilon M_1} (x_i-\nu_i)\, u_i'(\nu_i) + \tfrac{1}{2} \sum_{i\varepsilon M_1} (x_i-\nu_i)^2\, u_i''(\nu_i) + \sum_{i\varepsilon M_2} u_i(\tau) + \sum_{i\varepsilon M_2} (x_i-\tau)\, u_i'(\tau) + \tfrac{1}{2} \sum_{i\varepsilon M_2} (x_i-\tau)^2\, u_i''(\nu_i) .$$

In accordance to (2)

$$(8) \qquad y = \sum_{i\varepsilon M} x_i = \sum_{i\varepsilon M_1} \nu_i + \sum_{i\varepsilon M_2} \tau$$

and any variable, for example x_k, $k \varepsilon M_1$, can be expressed as a function of other (n-1) variables:

$$(9) \qquad x_k = \sum_{i\varepsilon M_1} \nu_i + \sum_{i\varepsilon M_2} \tau - \sum_{\substack{i\varepsilon M_1 \\ i\neq k}} x_i - \sum_{i\varepsilon M_2} x_i \quad .$$

Replacing (7) into (6) by taking into consideration (9), one can obtain

$$\begin{aligned}
\Delta = & \sum_{i\varepsilon M_1} u_i(\nu_i) + \sum_{i\varepsilon M_2} u_i(\tau) - \sum_{i\varepsilon M_1} u_i(\nu_i) - \sum_{i\varepsilon M_2} u_i(\tau) \\
& - \sum_{\substack{i\varepsilon M_1 \\ i\neq k}} (x_i - \nu_i)\, u_i'(\nu_i) - \Bigg(\sum_{i\varepsilon M_1} \nu_i - \nu_k + \sum_{i\varepsilon M_2} \tau \\
(10) \qquad & - \sum_{\substack{i\varepsilon M_1 \\ i\neq k}} x_i - \sum_{i\varepsilon M_2} x_i \Bigg) u_i'(\nu_i) - \sum_{i\varepsilon M_2} (x_i - \tau)\, u_i'(\tau) \\
& - \tfrac{1}{2} \sum_{i\varepsilon M_1} (x_i - \nu_i)^2\, u_i''(\nu_i) - \tfrac{1}{2} \sum_{i\varepsilon M_2} (x_i - \tau)^2\, u_i''(\nu_i) \quad .
\end{aligned}$$

Since for any concave function $u_i''(x_i) \leq 0$, the final form of the equation (10) is

$$\begin{aligned}
\Delta = & \sum_{i\varepsilon M_2} (x_i - \tau) \left[u_i'(\nu_i) - u_i'(\tau) \right] + \tfrac{1}{2} \sum_{i\varepsilon M_1} (x_i - \nu_i)^2 \left| u_i''(\nu_i) \right| \\
& + \tfrac{1}{2} \sum_{i\varepsilon M_2} (x_i - \tau)^2 \left| u_i''(\tau) \right| \quad .
\end{aligned}$$

It is obvious that as by the condition of the theorem $u_i'(\tau) \leq \lambda_1 = u_i'(x_i)$ for any x_i, $i \in M_2$ the difference $\Delta \geq 0$. Therefore, the theorem is proved.

Computational Procedures

Following the theorem three types of computational procedures could be suggested.

I. Graphical solution of the problem

For solving the problem it is sufficient to draw all the equations $u_i'(x_i) = \lambda$, $i = \overline{1,n}$ for any $\lambda \leq \max_i u_i'(\tau)$ as is shown in Fig. 1.*) After the drawing has been done, for every value of λ the sum $\sum_{i=1}^{n} x_i(\lambda)$ is computed and hence, in accordance to (2) the function $\lambda(y)$ is obtained.

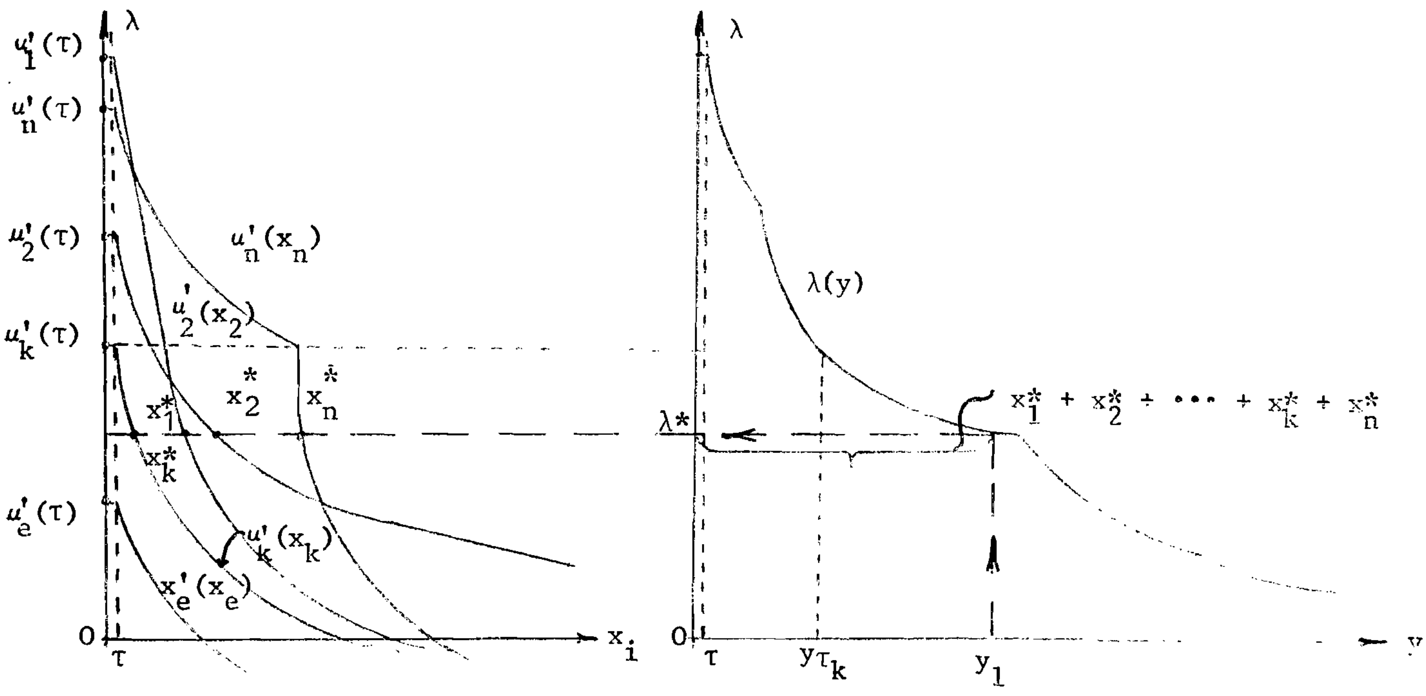


Fig. 1

*) In the case when the constraint (2) is in the form $\sum_{i=1}^{n} x_i \leq y$, $x_i \geq 0$, it is sufficient to consider only λ in the interval $0 \leq \lambda \leq \max_i u_i'(\tau)$, because for $\lambda=0$ the function $\sum_{i=1}^{n} u_i(x_i)$ reaches its absolute maximum (maximum without constraints).

Having $\lambda(y)$ the calculation of the optimal vector is very simple. As is shown in Fig. 1 for a given $y = y_1$ the optimum value of λ is λ^* and the optimal vector $x^* = (x_1^*, x_2^*, \ldots, x_e^* = 0, x_k^*, x_n^*)$.

The graphical solution enables us to find the solution "at once" when y is changed, i.e. to find parametrical solutions.

II. Analytical solution of the problem in closed form

Using Fig. 1 the following equation can be written

$$(11) \qquad x_i(\lambda) = \begin{cases} \varphi_i(\lambda) & , \text{ if } u_i'(\tau) - \lambda > 0 \\ \tau & , \text{ if } u_i'(\tau) - \lambda \leq 0 \end{cases} , \quad i = \overline{1,n} ,$$

or in more compact form

$$(12) \qquad x_i(\lambda) = \varphi_i(\lambda)\ 1\left[u_i'(\tau) - \lambda\right] + \tau .\ 1\left[\delta + \lambda - u_i'(\tau)\right] , \quad i = \overline{1,n}$$

$$\text{where } 1(\cdot) = \begin{cases} 1, & \text{if } (\cdot) > 0 \\ 0, & \text{otherwise} \end{cases}$$

δ is any positive number and it shows that when $\lambda - u_i'(\tau) = 0$, then $x_i(\lambda)$ is equal to τ.

Let us denote with y_{τ_k} the value of y when $\lambda = u_k'(\tau)$, i.e.

$$(13) \qquad y_{\tau_k} = \sum_{i=1}^{n} x_i(\lambda)\Big|_{\lambda = u_k'(\tau)} = \sum_{i=1}^{n} \Big\{\varphi_i(\lambda) . 1\left[u_i'(\tau) - \lambda\right] + \tau .\ 1\left[\delta + \lambda - u_i'(\tau)\right]\Big\} .$$

It is obvious from Fig. 1 that when y is changed, i.e. $y \geq y_{\tau_k}$ or $y < y_{\tau_k}$, the variable $x_k(\lambda)$ is

(14) $$x_k(\lambda) = \begin{cases} \varphi_k(\lambda) , & \text{if } y - y_{\tau_k} > 0 \\ \tau , & \text{if } y - y_{\tau_k} \leq 0 \end{cases} , \quad k = \overline{1,n} ,$$

or in shortened form

(15) $$x_k(\lambda) = \varphi_k(\lambda) \; 1(y - y_{\tau_k}) + \tau . \; 1(\delta + y_{\tau_k} - y) .$$

According to the constraint (2), namely $\sum_{i=1}^{n} x_k(\lambda) = y$,

(16) $$\sum_{k=1}^{n} \varphi_k(\lambda) . 1(y - y_{\tau_k}) + \tau . \; 1(\delta + y_{\tau_k} - y) = y .$$

In this equation, representing λ as an implicit function of y, only the variable λ is unknown and by solving it the optimum value of $\lambda = \lambda^*$ can be calculated.

Having λ^* one can obtain the optimum value of all the variables x_k replacing $\lambda = \lambda^*$ in (15), i.e.

(17) $$x_k^*(\lambda^*) = \varphi_k(\lambda^*) \; 1(y - y_{\tau_k}) + \tau . \; 1(\delta + y_{\tau_k} - y) , \quad k = \overline{1,n} .$$

The whole procedure for determining the optimal vector x^* can be summarized as follows:

1. Choose τ
2. Calculate $u_k'(\tau)$, $k = \overline{1,n}$
3. Calculate y_{τ_k} , $k = \overline{1,n}$ from (13)
4. Solve the equation (16) in respect to λ
5. Replace λ^* in (17). The obtained values $x_k^*(\lambda^*)$ are components of the optimal vector x^*.

III. Seeking method for the analytical solving of the problem

Sometimes great computational difficulties will be met when one solves the equation (16). To overcome these difficulties, the following method can be proposed.

The method is based on the so-called "golden cut" [2] which usually is applied for seeking the maximum of a unimodal function.

Let us denote the upper and the lower boundaries of λ with G and A respectively (Fig. 2). The upper boundary $G = \max_i u_i'(\tau)$, while the lower boundary has to be a greater number approaching $(-\infty)$. In the case when the constraint (2) is in the form $\sum_{i=1}^{n} x_i \leq y$, $A = \tau$.

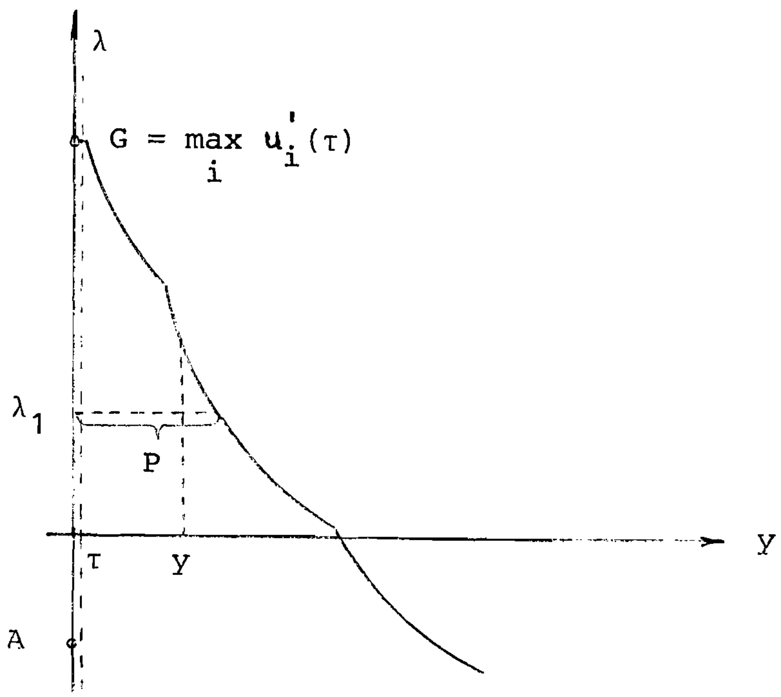


Fig. 2

Under these conditions the following procedure can be applied:

1. Determine $\lambda_1 = A + \frac{G - A}{1.61803}$

2. Determine $x_i = \begin{cases} \varphi_i(\lambda) & , \text{ if } u_i'(\tau) - \lambda > 0 \\ \tau & , \text{ if } u_i'(\tau) - \lambda \leq 0 \end{cases}$, $i = \overline{1,n}$

3. Denote $\sum_{i=1}^{n} x_i = P$. If $P = y$, then $x_i = x_i^*$, $i = \overline{1,n}$ and the procedure is terminated.

If $P > y$ (as is shown in Fig. 2), then $A = \lambda_1$ and we have to return to the first item of the procedure.

If $P < y$, then $G = \lambda_1$ and we return to the first item.

It has to be mentioned that using the similar procedures the following optimization problem can be solved:

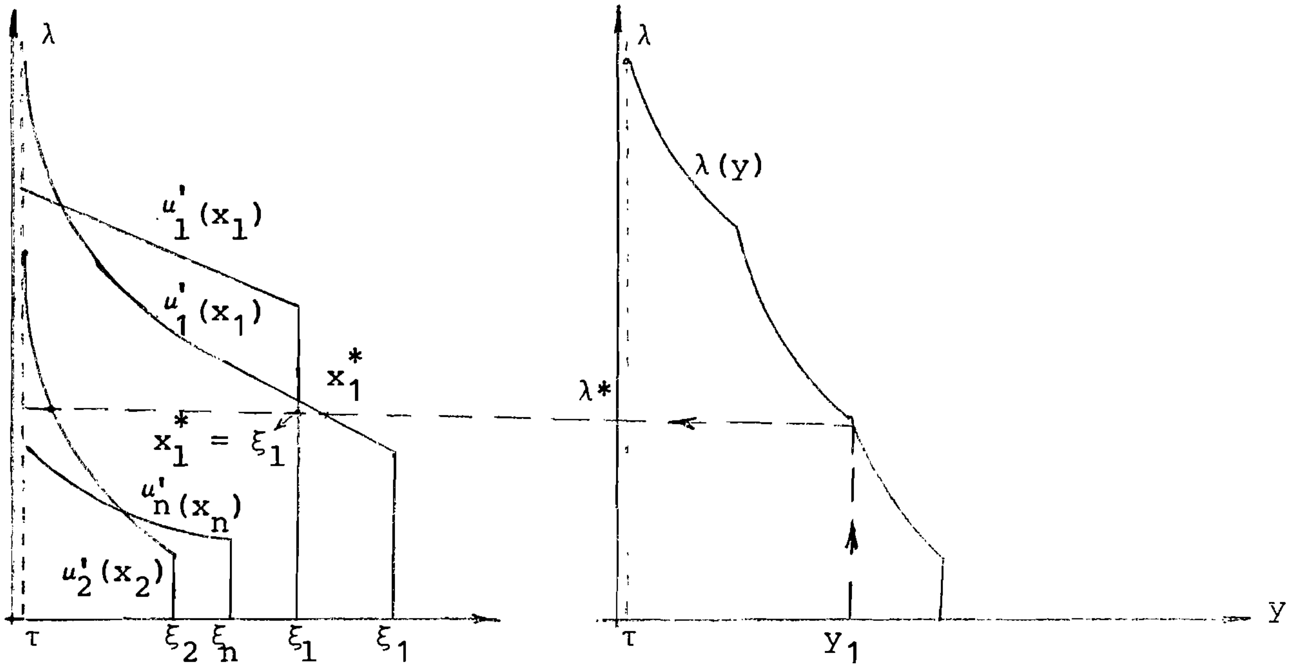
$$\max \sum_{i=1}^{n} u_i(x_i)$$

subject to

$$\sum_{i=1}^{n} x_i \leq y$$

$$\tau \leq x_i \leq \xi_i \quad , \qquad i = \overline{1,n} \quad .$$

For this problem only a graphical solution is shown in Fig. 3.



Example. Using a very simple example an analytical solution of the problem in closed form will be shown.

$$\max \left[50\, x_1 - 2x_1^2 + 100x_2 - x_2^2 + 200 \sqrt{x} \right]$$

subject to

$$\sum_{i=1}^{3} x_i = 25 \quad ; \quad x_i \geq 0 \quad , \quad i = 1,2,3$$

1. The first step of the algorithms is the choosing of τ. Let $\tau = 0.001$.

2. Calculation of $u_k'(\tau)$, $k = 1,2,3$.

$u_1'(\tau) \doteq 49.996$; $u_2'(\tau) = 99.998$; $u_3'(\tau) = 3333.33$.

3. Calculation of y_{τ_k}, $k = 1,2,3$ from (13)

$y_{\tau_1} = 29.002$; $y_{\tau_2} = 1.003$; $y_{\tau_3} = 0.003$.

4. After replacing y_{τ_k}, $k = 1,2,3$ in (16) we obtain

$$\frac{100 - \lambda}{2} + \frac{10000}{\lambda^2} = 25 \quad .$$

Solving this equation the optimum value of λ is obtained, $\lambda^* = 56.308$.

5. The components of optimum vector x are obtained by (17) after replacing $\lambda = \lambda^*$, i.e.

$$x_1^* = 0; \quad x_2^* = 21.85; \quad x_3^* = 3.15 \quad .$$

References

[1] Rozanov, Yu. A. Optimum Fund Distribution, IIASA working paper WP-74-61, November 1974.

[2] Wilde, D.J. Optimum-Seeking Methods, 1964.