JOURNAL
OF
APPLIED
QUANTITATIVE
METHODS

SOLVING NONLINEAR OPTIMIZATION PROBLEMS BY MEANS OF THE NETWORK PROGRAMMING METHOD

Vladimir N. BURKOV

Prof., Institute of Control Sciences of V.A. Trapeznikov,
Russian Academy of Sciences,
Moscow, Russia

E-mail: vlab17@bk.ru

Irina V. BURKOVA

PhD Candidate, Institute of Control Sciences of V.A. Trapeznikov,
Russian Academy of Sciences,
Moscow, Russia

E-mail: irbur27@gmail.com

Abstract: *We suggest a new approach to solve discrete optimization problems, based on the possibility of presenting a function as a superposition of simpler functions. Such a superposition can be easily represented in the form of a network for which the inputs correspond to variables, intermediate nodes – to functions entering the superposition, and in the final node the function is calculated. Due to such representation the method has been called the method of network programming (in particular, dichotomic). The network programming method is applied for solving nonlinear optimization problems. The concept of a dual problem is implemented. It is proved that the dual problem is a convex programming problem. Necessary and sufficient optimality conditions for a dual problem of integer linear programming are developed.*

Key words: *network programming; nonlinear optimization; dual problem; integer linear programming*

1. Introduction

Problems of nonlinear optimization (in particular, discrete optimization) refer to the class of so-called NP-difficult problems for which no effective methods of exact solution do exist. Some general approaches are available, among others the branch and bounds method and the method of dynamic programming [1]. Unfortunately, the dynamic programming method is applicable only to a narrow class of problems. The efficiency of the branch and bounds method depends essentially on accuracy of the upper and lower estimates (bounds).

To assess those estimates the method of multipliers of Lagrange [1] is developed. These methods are known from the past 60-s, and since then more than they have not been improved significantly.

In 2004 V.N.Burkov and I.V.Burkova suggested a new approach to solve discrete optimization problems, based on the possibility of presenting a function as a superposition of simpler functions. Such a superposition can be easily represented in the form of a network for which the inputs correspond to variables, intermediate nodes – to functions entering the superposition, and in the final node the function is calculated. Due to such representation the method has been called the method of network programming [2] (in particular, dichotomic). This method is applicable to cases when the goal function and restriction functions obtain identical network structure. For such cases network node optimization problems, simpler than the pregiven ones, are solved. The problems' solution for the final node presents the upper (or lower) estimates for the given problem. For the case when the network structure is a tree, the solution becomes an exact one. The Bellman's dynamic programming method for which the network structure comprises tree branches, becomes thus a particular case of the more generalized proposed approach. A variety of problems for which the dynamic programming method is inapplicable, have been solved by the network programming method.

In the present paper the network programming method [2] is applied to nonlinear programming problems. The concept of a dual problem, for which one of the feasible (but usually non-optimal) solutions is obtained, is suggested by means of multipliers of Lagrange. It is proved that the dual problem is a convex programming problem. Necessary and sufficient optimality conditions for a dual problem of integer linear programming are developed.

2. The Network Form of a Nonlinear Programming Problem

Let's consider a problem of nonlinear programming - to determine $x=\{x_i,\ i=\overline{1,\ n}\}$, satisfying

$$f(x)\to \max \tag{1}$$

subject to

$$\varphi_j(x)\le b_j, \quad j=\overline{1,m}, \tag{2}$$

$$x\in X_{m+1}. \tag{3}$$

On Figure 1 the network representation of restrictions (2-3) is given. Here X_j denotes the j -th restriction (2), $j=\overline{1,m}$.

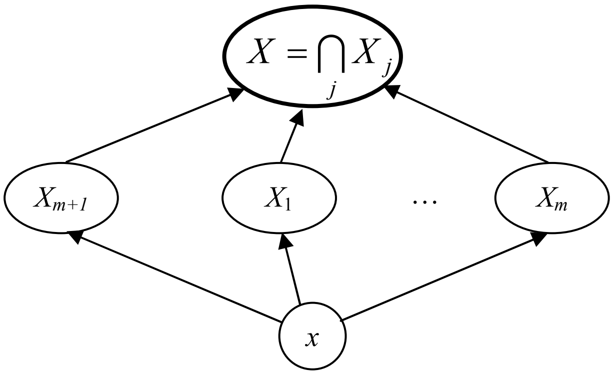

Figure 1. Network representation of restrictions

In order to apply the network programming method we have to represent the goal function with the same network structure. For this purpose we will present *f(x)* in the form

$$f(x) = \sum_{j=1}^{m} h_j(x) + h_{m+1}(x), \tag{4}$$

where $h_i(x)$ stands for functions which deliver solutions for the below problems (5-6).

In each vertex of the network structure several optimization sub-problems with one restriction are solved. The first *m* sub-problems are as follows:

$$\begin{aligned} &\max h_j(x), \\ &\varphi_j(x) \le b_j. \end{aligned} \tag{5}$$

while the *(m+1)* -th sub-problem looks as follows:

$$\max_{x \in X} h_{m+1}(x) = \max_{x \in X_{m+1}} \left[f(x) - \sum_{j=1}^{m} h_i(x) \right]. \tag{6}$$

Denote $F_i(h)$ the value of the goal function for the optimal solution of the *j* -th sub-problem.

Theorem 1. Linear model

$$F(h) = \sum_{j=1}^{m} F_j(h_j) + F_{m+1}(h) \tag{7}$$

delivers the upper estimate for a pre-given problem.

Proof. All feasible solutions (1-3) are feasible for all sub-problems (5-6), and any feasible solution x satisfies

$$\sum_{j=1}^{m+1} h_j(x) = f(x)$$

.

Therefore *F(h)* ≥ *f(x)* for any feasible x.

3. The Dual Problem

It is obvious to suggest the problem of determining functions $h_i(x)$, $j = \overline{1, m}$, which minimize the upper estimate (7). This problem is, in essence, a generalized dual problem for the initial problem of nonlinear programming. The reasons for this are as follows. First, as shown below (see Example 1), one of the feasible solutions of the generalized dual problem is a minimax of function of Lagrange. Note that determining the minimax Lagrange function is often called the dual one for the problem of nonlinear programming. Second, for a problem of linear programming without an integer solution the generalized dual problem is a usual dual problem of linear programming (see Section 4).

Theorem 2. Function *F(h)* is a convex one.

Proof. Let $h_1(x)$ and $h_2(x)$ be two solutions of a dual problem. Consider the solution

$$h(x) = \alpha h^1 + (1-\alpha) h^2, \quad 0 \le \alpha \le 1.$$

We obtain

$$F(h) = F\left[\alpha h^1 + (1-\alpha)h^2\right] = \max_{x\in X}\left[\alpha\left(f(x) - \sum_{j=1}^{m} h_j^1(x)\right) + (1-\alpha)\left(f(x) - \sum_{j=1}^{m} h_j^2(x)\right)\right] +$$

$$+\sum_{j=1}^{m} \max_{\varphi_j(x)\le b_j}\left[\alpha h_j^1(x) + (1-\alpha)h_j^2(x)\right] \le \alpha\left[\max_{x\in X}\left(f(x) - \sum_{j=1}^{m} h_j^1(x)\right) + \sum_{j=1}^{m}\max_{\varphi_j(x)\le b_j} h_j^1(x)\right] +$$

$$+(1-\alpha)\left(\max_{x\in X}\left(f(x) - \sum_{j=1}^{m} h^2(x)\right) + \sum_{j=1}^{m}\max_{\varphi_j(x)\le b_j} h_j^2(x)\right) = \alpha F(h^1) + (1-\alpha)F(h^2).$$

The inequality stems from the evident reason that the maximum of the sum is less or equal to the sum of maxima.

Thus, the dual problem is a convex programming problem.

Example 1. Consider one of the feasible solutions of the dual problem, namely, $h_j(x) = \lambda_j \varphi_j(x),\ j = \overline{1,m}$. The first *m* sub-problems are as follows:

$$\lambda_j \varphi_j(x) \to \max$$

subject to

$$\varphi_j(x) \le b_j.$$

Evidently $F_j(h_j) \le \lambda_j b_j, \quad j = \overline{1,m}$. By means of this assertion together with (7), we finally obtain

$$F(\lambda) \le \max_{x\in X_{m+1}}\left(f(x) - \sum_{j=1}^{m}\lambda_j\left(\varphi_j(x) - b_j\right)\right) = \max_{x\in X_{m+1}} L(\lambda, x). \tag{8}$$

Maximizing the right part (8) on λ is nothing but the method of multipliers of Lagrange. Thus, the method of multipliers of Lagrange provides a feasible solution of the dual problem (which, generally speaking, may be not an optimal one).

4. Upon One Integer Linear Programming Problem

Consider an integer linear programming problem as follows: determine an integer nonnegative vector x, to maximize

$$C(x) = \sum_{i=1}^{n} c_i x_i \tag{9}$$

subject to

$$\sum_{i=1}^{n} a_{ij} x_i \le b_j, \quad j = \overline{1, m+1}. \tag{10}$$

Take the last restriction in (10) as the set X_{m+1}. Sub-divide each value $c_i, i=\overline{1,m}$, on *m* partial values s_{ij} as follows:

$$s_{i;m+1} = c_i - \sum_{j=1}^{m} s_{ij}, \quad i = \overline{1,n}. \tag{11}$$

Solve $(m+1)$ sub-problems as follows: determine an integer nonnegative vector *x*, to maximize

$$S_j(x) = \sum_i s_{ij} x_i . \tag{12}$$

subject to

$$\sum_{i=1}^{n} a_{ij} x_i \le b_j . \tag{13}$$

Denote by $F_j(s)$ the value $S_j(x)$ providing the optimal solution for the *j* -th subproblem. According to Theorem 1

$$F(s) = \sum_{j=1}^{m} F_j(s_j) + F_{m+1}(s) \tag{14}$$

is an upper estimate for $C(x)$:

$$F(s) \ge C(x).$$

The dual problem: determine $\{s_{ij},\ i=\overline{1,n}, j=\overline{1,m}\}$, minimizing (14). Note that cancelling the requirement of integrality results in transforming problem (14) to a common dual linear programming problem [3].

To prove this accession consider problem (9-10) without the integrality requirement. In this case the estimation problems are easily solved, namely

$$F_j(s_j) = b_j \max_i \frac{s_{ij}}{a_{ij}}.$$

Denote

$$y_j = \max_i \frac{s_{ij}}{a_{ij}}, \quad j = \overline{1,m+1}.$$

Thus, the upper estimate for the objective of the initial problem looks as follows:

$$\Phi(y) = \sum_j y_j b_j . \tag{15}$$

Since $a_{ij}\, y_j \ge s_{ij}$, relation (11) transfers to

$$\sum_j a_{ij} y_j \ge c_i, \quad i = \overline{1,n}. \tag{16}$$

The dual problem is to minimize (15) subject to (16). This is a common dual linear programming problem.

Set $s_{ij} = \lambda_j \alpha_{ij},\ i=\overline{1,n}, j=\overline{1,m}$. As outlined above, the problem boils down to the method of multipliers of Lagrange as follows: determine vector λ, minimizing

$$\max_{x \in X_{m+1}} \left(\sum_i c_i x_i - \sum_{j=1}^{m} \lambda_j \left(\sum_i a_{ij} - b_j \right) \right). \tag{17}$$

Consider necessary and sufficient conditions to obtain the optimal solution of the dual problem. Let *s* be a feasible solution. Denote P_j (s_j) the set of optimal solutions for $(m+1)$ sub-problems (12-13), $j=\overline{1,m+1}$.

Theorem 3. The necessary and sufficient condition to obtain the optimal solution *s* is the inability to solve inequality

$$\sum_j \max_{x\in P_j(s_j)} \sum_i y_{ij} x_i < 0 \qquad (18)$$

subject to

$$\sum_{j=1}^{m+1} y_{ij} = 0, \quad i=\overline{1,n}. \qquad (19)$$

Proof. Denote by y_{ij} small increments of s_{ij}. We will prove that relations (19) stem from (11). Indeed, it boils down from (11) that

$$\sum_{j=1}^{m+1}\left(y_{ij}+s_{ij}\right)=c_i \quad \text{and} \quad \sum_{i=1}^{m+1} s_{ij}=c_i$$

hold. The latter provides (19). The increment of value F_j (s_j) is, obviously, equal

$$\Delta F_j = \max_{x\in P_j(s_j)} \sum_i y_{ij} x_i,$$

while the total increment satisfies

$$\Delta F = \sum_j \Delta F_j.$$

Since *s* is the optimal solution, ΔF cannot be negative. ■

Numerical Example 2. $x_i = 0, 1$; $i = \overline{1,4}$.

$$10x_1+8x_2+6x_3+7x_4 \to \max, \qquad (20)$$

$$6x_1+3x_2+2x_3+5x_4 \le 11, \qquad (21)$$

$$3x_1+5x_2+6x_3+3x_4 \le 11. \qquad (22)$$

Apply the method of multipliers of Lagrange, i.e. determine the minimum of λ functions

$$11\lambda + \max_{x\in X_2}\left[(10-6\lambda)x_1+(8-3\lambda)x_2+(6-2\lambda)x_3+(7-5\lambda)x_4\right],$$

where X_2 is determined by (22). With pre-set λ this is a one-dimensional knapsack problem. In case when the dependence of the right part of b_2 (see restriction (22)) from *n* is unknown, this problem turns to be NP-difficult [4]. However in practice, b_2 either does not depend on *n*, or is a linear function of *n*. In such cases, for integer parameters, the problem is efficiently solved by means of either dynamic or dichotomic programming. The determined optimal value $\lambda_0 = 1\tfrac{2}{9}$, with the upper estimate $F_0 = 21\tfrac{1}{3}$. This level λ_0 corresponds to the following values s_{ij}, $i=\overline{1,4}$, $j=1,2$:

$$\begin{array}{llll} s_{11}=\lambda_0 a_{11}=7\tfrac{1}{3}; & s_{21}=\lambda_0 a_{21}=3\tfrac{2}{3}; & s_{31}=\lambda_0 a_{31}=2\tfrac{4}{9}; & s_{41}=\lambda_0 a_{41}=6\tfrac{1}{9}; \\ s_{12}=c_1-s_{11}=2\tfrac{2}{3}; & s_{22}=4\tfrac{1}{3}; & s_{32}=3\tfrac{5}{9}; & s_{42}=\tfrac{8}{9}. \end{array} \qquad (23)$$

Let's apply the network programming method.

Step 1. Determine necessary optimality conditions for solution (23). Consider:

$$P_1(s_1) = \{(1,1,1,0);\quad (1,0,0,1)\};$$
$$P_2(s_2) = \{(1,1,0,1);\quad (0,1,1,0)\}.$$

Since $y_{i1} + y_{i2} = 0$, denote $y_i = y_{i1} = -y_{i2}$. In this case relations (18-19) can be represented as

$$\max(y_1 + y_2 + y_3; y_1 + y_4) < \min(y_1 + y_2 + y_4; y_2 + y_3).$$

One of the solutions for those relations is as follows:

$$y_1 = -\varepsilon;\quad y_2 = \varepsilon;\quad y_3 = -\varepsilon;\quad y_4 = 0;\quad \varepsilon > 0.$$

Set $\varepsilon = 5/6$; since this value results in a new solution of the second sub-problem. We obtain:

$$s_{11} = 6\,1/2;\quad s_{21} = 4\,1/2;\quad s_{31} = 1\,11/18;\quad s_{41} = 6\,1/9;$$
$$s_{12} = 3\,1/2;\quad s_{22} = 3\,1/2;\quad s_{32} = 4\,7/18;\quad s_{42} = 8/9;$$
$$P_1(s_1) = \{(1,0,0,1);\ (1,1,1,0)\};\quad F_1 = 12\,11/18;$$
$$P_2(s_2) = \{(1,1,0,1);\ (0,1,1,0);\ (1,0,1,0)\};\quad F_2 = 7\,8/9;\qquad F = 20\,1/2.$$

Step 2. Consider optimality conditions

$$\max(y_1 + y_2 + y_3; y_1 + y_4) < \min(y_1 + y_2 + y_4; y_2 + y_3; y_1 + y_3).$$

It can be well-recognized that this inequality has no feasible solutions. Indeed, condition y1 + y2 + y3 < y1 + y2 + y4 results in $y_3 < y_4$, while condition $y_2 + y_4 < y_2 + y_3$ leads to a contradictive $y_4 < y_3$. Hence, the optimal solution of the dual problem is obtained. The determined upper estimate may be used in the branch and bounds method. Start branching with variable x_1. If $x_1 = 1$ then the solution of the corresponding dual problem results in the same estimate $F(x_1=1) = 20^1/_2$. In case $x_1 = 0$ the obtained estimate $F(x_1=0) = 14$. Choose value $x_1 = 1$ and undertake branching for variable x_2. $x_2 = 1$ results in a feasible estimate $F(x_1=1, x_2=1) = 18$. $x_2 = 0$ results in another feasible estimate $F(x_1=1, x_2=0) = 17$. Thus, the optimal solution is $x_1 = 1, x_2 = 1, x_3 = 0, x_4 = 0, C_{max} = 18$.

5. Conclusions

The suggested approach provides a generalized method to determine estimates for a broad class of nonlinear programming problems. This approach enables using new algorithms to solve a variety of problems, with the computing complexity being less, than that when using classical algorithms (the knapsack problem [3], the maximal flow problem [5], the "stones" problem [3], etc.). Further research has to be undertaken to estimate the computing complexity of the network programming method for various problems of nonlinear programming.

References

1. Burkov, V.N., Zalozhnev, A.J. and Novikov, D.A. **Graph Theory in Managing Organizational Systems,** Sinteg, Moscow, 2001 (in Russian)
2. Burkov, V.N. and Burkova, I.V. **Network programming method,** Management Problems, 3, 2005, pp. 23-29 (in Russian)
3. Burkov, V.N. and Burkova, I.V. **Method of dichotomizing programming,** Institute of Control Sciences, the Russian Academy of Sciences, 2004, pp. 57-75 (in Russian)

4. Gary, M. and Johnson, D. **Computers and Difficultly-Solved Problems,** Mir, Moscow, 1982 (in Russian)
5. Burkov, V.N. (Ed.), Mathematical backgrounds of project management, Manual, Visshaya Shkola, Moscow, 2005, pp. 312-336 (in Russian)