RUTCOR RESEARCH REPORT

Maximization of a Strongly Unimodal Multivariate Discrete Distribution

Mine Subasi[a] Ersoy Subasi[b] András Prékopa[c]

RRR 12-2009, July 2009

RUTCOR
Rutgers Center for
Operations Research
Rutgers University
640 Bartholomew Road
Piscataway, New Jersey
08854-8003
Telephone: 732-445-3804
Telefax: 732-445-5472
Email: rrr@rutcor.rutgers.edu
http://rutcor.rutgers.edu/∼rrr

[a]RUTCOR, Rutgers Center for Operations Research, 640 Bartholomew Road Piscataway, NJ 08854-8003, USA. Email: msub@rutcor.rutgers.edu
[b]RUTCOR, Rutgers Center for Operations Research, 640 Bartholomew Road Piscataway, NJ 08854-8003, USA. Email: esub@rutcor.rutgers.edu
[c]RUTCOR, Rutgers Center for Operations Research, 640 Bartholomew Road Piscataway, NJ 08854-8003, USA. Email: prekopa@rutcor.rutgers.edu

Rutcor Research Report
RRR 12-2009, July 2009

Maximization of a Strongly Unimodal Multivariate Discrete Distribution

Mine Subasi Ersoy Subasi András Prékopa

Abstract. A dual type linear programming algorithm is presented for locating the maximum of a strongly unimodal multivariate discrete distribution.

1 Introduction

A probability measure P, defined on $\mathbb{R}^n$, is said to be logconcave if for every pair of nonempty convex sets $A, B \subset \mathbb{R}^n$ (any convex set is Borel measurable) and we have the inequality

$$P(\lambda A + (1-\lambda)B) \geq [P(A)]^\lambda [P(B)]^{(1-\lambda)},$$

where the + sign refers to Minkowski addition of sets, i.e.,

$$\lambda A + (1-\lambda)B = \{\lambda \mathbf{x} + (1-\lambda)\mathbf{y} | \mathbf{x} \in A, \mathbf{y} \in B\}.$$

The above notion generalizes in a natural way to nonnegative valued measures. In this case we require the logconcavity inequality to hold for finite $P(A), P(B)$. The notion of a logconcave probability measure was introduced in [8, 9].

In 1912 Fekete [4] introduced the notion of an r-times positive sequence. The sequence of nonnegative elements $\ldots, a_{-2}, a_{-1}, a_0, \ldots$ is said to be *r-times positive* if the matrix

$$A = \begin{bmatrix} \ddots & \ddots & \ddots & & \\ \ddots & a_0 & a_1 & a_2 & \\ \ddots & a_{-1} & a_0 & a_1 & \ddots \\ & a_{-2} & a_{-1} & a_0 & \ddots \\ & & \ddots & \ddots & \ddots \end{bmatrix}$$

has no negative minor of order smaller than or equal to r.

Twice-positive sequences are those for which we have

$$\begin{vmatrix} a_i & a_j \\ a_{i-t} & a_{j-t} \end{vmatrix} = a_i a_{j-t} - a_j a_{i-t} \geq 0. \tag{1.1}$$

for every $i < j$ and $t \geq 1$. This holds if and only if

$$a_i^2 \geq a_{i-1} a_{i+1} \ .$$

Fekete [4] also proved that the convolution of two r-times positive sequences is r-times positive. Twice-positive sequences are also called logconcave sequences. For this, Fekete's theorem states that the convolution of two logconcave sequences is logconcave.

A discrete probability distribution, defined on the real line, is said to be logconcave if the corresponding probability function is logconcave. While continuous unimodal or convex functions enjoy a number of useful properties, many of them do not carry over the discrete case (see, e.g., [10]). For example, the convolution of two logconcave distributions on $\mathbb{Z}^n$, the set of lattice points in the space, is no longer logconcave in general, if $n \geq 2$.

Favati and Tardella [3] introduced a notion of integer convexity. They analyzed some connections between the convexity of a continuous function (on $\mathbb{R}^n$) and integer convexity of its restriction to $\mathbb{Z}^n$. They also presented a polynomial-time algorithm to find the minimum of a submodular integrally convex function. A further paper in this respect is due to Murota [6]. He developed a theory of discrete convex analysis for integer valued functions defined on integer lattice points.

A classical paper on discrete unimodality is due to Barndorff-Nielsen [1] in which a notion of strong unimodality was introduced. Following Barndorff-Nielsen [1] a discrete probability function $p(\mathbf{z}), \mathbf{z} \in \mathbb{Z}^n$ is called strongly unimodal if there exists a convex function $f(\mathbf{x}),\ \mathbf{x} \in \mathbb{R}^n$ such that $f(\mathbf{x}) = -\log p(\mathbf{x})$ if $\mathbf{x} \in \mathbb{Z}^n$. If $p(\mathbf{z}) = 0$, then by definition $f(\mathbf{z}) = \infty$. This notion is not a direct generalization of that of the one-dimensional case, i.e., of formula 1.1. However in case of $n = 1$ the two notions are the same (see, e.g., [10]). It is trivial that if p is strongly unimodal, then it is logconcave. The joint probability function of a finite number of mutually independent discrete random variables, where each has a logconcave probability function is strongly unimodal.

Pedersen [7] presented sufficient conditions for a bivariate discrete distribution to be strongly unimodal. He also proved that the trinomial probability function is logconcave and the convolution of any finite number of these distributions with possibly different parameter sets is also logconcave.

Subasi et al. [13] gave sufficient conditions that ensure the strong unimodality of a multivariate discrete distribution. In view of this they presented a subdivision of $\mathbb{R}^n$ into non-overlapping convex polyhedra such that $f(\mathbf{x})$ coincides with the values of $\log p(\mathbf{x}),\ \mathbf{x} \in \mathbb{Z}^n$, is linear on each of them. On each subdividing simplex $f(\mathbf{x})$ is defined by the equation of the hyperplane determined by the values of $-\log p(\mathbf{x})$ at the vertices and the convexity of $f(\mathbf{x})$ on any two neighboring simplices (two simplices having a common facet) is ensured. The resulting function $f(\mathbf{x})$ is convex on the entire space. In the same paper the authors also proved the strong unimodality of the negative multinomial distribution, the multivariate hypergeometric distribution, the multivariate negative hypergeometric distribution and the Dirichlet (or Beta) compound multinomial distribution. These theoretical investigations lead to some practical suggestion on how to find the maximum of a strongly unimodal multivariate discrete distribution.

A function $f(\mathbf{z}), \mathbf{z} \in \mathbb{R}^n$ is said to be polyhedral (simplicial) on the bounded convex polyhedron $K \subseteq \mathbb{R}^n$ if there exists a subdivision of K into n-dimensional convex polyhedra (simplices), with pairwise disjoint interiors such that f is continuous on K and linear on each subdividing polyhedron (simplex). Prékopa and Li [11] presented a dual method to solve a linearly constrained optimization problem with convex, polyhedral objective function, along with a fast bounding technique, for the optimum value. Any $f(x)$, defined by the use of a strongly unimodal probability function $p(x)$, is a simplicial function and can be used in the above-mentioned methodology. In an earlier paper [**?**] Prékopa developed a dual type method for the solution of a one-stage stochastic programming problems. The method was improved and implemented in [2].

The notion of discrete unimodality is of interest, for example, in connection with statistical physics where a typical problem is to find the maximum of a unimodal function. In this paper we use the results of the paper by Prékopa and Li [11] and present a dual type algorithm to locate the maximum of a strongly unimodal multivariate discrete distribution. In what follows we present our results in terms of probability functions. They generalize in a straightforward manner for more general discrete unimodal functions.

2 A Dual Type Algorithm

The problem of interest is to find the maximum of a strongly unimodal probability function $p(\mathbf{x}),\ \mathbf{x} \in \mathbb{Z}^n$ which is the same as the minimum value of the convex function $f(x)$ on $\mathbb{R}^n$ such that

$$f(x) = -\log p(\mathbf{x}) \ \text{ if } \ \mathbf{x} \in \mathbb{Z}^n$$

that we will be looking at.

In addition to the strongly unimodality of $p(\mathbf{x}),\ \mathbf{x} \in \mathbb{Z}^n$ we assume that there exists a subdivision of $\mathbb{R}^n$ into simplices with pairwise disjoint interiors such that all vertices of all simplices are elements of $\mathbb{Z}^n$ and a function $f(\mathbf{x}),\ \mathbf{x} \in \mathbb{R}^n$ that is linear on each of them, otherwise convex in $\mathbb{R}^n$ and $f(\mathbf{x}) = -\log p(\mathbf{x})$ if $\mathbf{x} \in \mathbb{Z}^n$.

Probability functions frequently take zero values on some points of $\mathbb{Z}^n$. If $p(\mathbf{x}) = 0$, then by definition $f(\mathbf{x}) = \infty$ and therefore any $\mathbf{x}$ with this property can be excluded from the optimization. We can also restrict the optimization to bounded sets. In fact, since

$$\sum_{\mathbf{x} \in \mathbb{Z}^n} p(\mathbf{x}) = 1$$

it follows that there exists a vector $\mathbf{b}$ such that the minimum of f is taken in the set $\{\mathbf{x} \mid |\mathbf{x}| \le \mathbf{b}\}$. Such a $\mathbf{b}$ can easily be found without the knowledge of the minimum of f, we simply take a $\mathbf{b}$ with large enough components. For simplicity we assume that the minimum of f is taken at some point of the set $\{x \mid 0 \le \mathbf{x} \le \mathbf{b}\}$ and such a $\mathbf{b}$ is known.

Probability functions sometimes are of the type where the nonzero probabilities fill up the lattice points of a simplex. An example is the multinomial distribution where $p(x_1, ..., x_n) > 0$ for the elements of the set

$$\{\mathbf{x} \in \mathbb{Z}^n \mid \mathbf{x} \ge 0,\ x_1 + ... + x_n \le n\}.$$

Taking this into account, we will be looking at the problems

$$\begin{array}{rl} \min & f(\mathbf{x}) \\ \text{subject to} & \\ & 0 \le \mathbf{x} \le \mathbf{b}, \quad \mathbf{x} \ \text{ integer} \end{array} \tag{2.1}$$

where $\mathbf{b}$ has positive integer components and

$$\begin{array}{rl} \min & f(\mathbf{x}) \\ \text{subject to} & \\ & x_1 + ... + x_n \le b \\ & \mathbf{x} \ge 0, \quad \mathbf{x} \ \text{ integer} \end{array} \tag{2.2}$$

where b is a positive integer.

Our assumption regarding the subdivision of $\mathbb{R}^n$ into simplices and the function $f(\mathbf{x}),\ \mathbf{x} \in \mathbb{R}^n$ carry over to the feasible sets in problems (2.1) and (2.2) in a natural way.

The integrality restriction of $\mathbf{x}$ can be removed from both problems (2.1) and (2.2). In fact, our algorithm not only produces an optimal $\mathbf{x}$ but also a subdividing simplex of which $\mathbf{x}$ is an element and at least one of the vertices of the resulting simplex also is an optimal solution.

First we consider problem (2.1). Let N be the number of lattice points of the set

$$\{\mathbf{x} \mid 0 \le \mathbf{x} \le \mathbf{b}\} .$$

Any function value $f(\mathbf{x}),\ 0 \le \mathbf{x} \le \mathbf{b}$ can be obtained by λ-representation, as the optimum value of a linear programming problem:

$$\begin{aligned} f(\mathbf{x}) \quad &= \min_{\lambda} \ \sum_{k=1}^{N} f(\mathbf{z}_k)\lambda_k \\ \text{subject to} \quad & \\ & \sum_{k=1}^{N} \mathbf{z}_k \lambda_k = \mathbf{x} \\ & \sum_{k=1}^{N} \lambda_k = 1 \\ & 0 \le \mathbf{x} \le \mathbf{b} \\ & \lambda \ge 0 \ . \end{aligned} \tag{2.3}$$

By the use of problem (2.3), problem (2.1) can be written in the following way:

$$\begin{aligned} \min_{\mathbf{x}} f(\mathbf{x}) \quad &= \min_{\lambda,\mathbf{x}} \ \sum_{k=1}^{N} f(\mathbf{z}_k)\lambda_k \\ \text{subject to} \quad & \\ & \sum_{k=1}^{N} \mathbf{z}_k \lambda_k = \mathbf{x} \\ & \sum_{k=1}^{N} \lambda_k = 1 \\ & \mathbf{x} \le \mathbf{b} \\ & \lambda \ge 0 \ , \quad \mathbf{x} \ge 0. \end{aligned} \tag{2.4}$$

Introducing slack variables we rewrite the problem as

$$
\begin{aligned}
\min_{\mathbf{x}} f(\mathbf{x}) \qquad & = \min_{\lambda,\mathbf{x}} \sum_{k=1}^{N} f(\mathbf{z}_k)\lambda_k \\
\text{subject to} \qquad & \\
& \sum_{k=1}^{N} \mathbf{z}_k\lambda_k - \mathbf{x} = 0 \\
& \sum_{k=1}^{N} \lambda_k = 1 \\
& \mathbf{x} + \mathbf{u} = \mathbf{b} \\
& \lambda \geq 0\ , \quad \mathbf{x} \geq 0\ , \quad \mathbf{u} \geq 0.
\end{aligned}
\tag{2.5}
$$

In order to construct an initial dual feasible basis to problem (2.5) we use the following theorem by Prékopa and Li [11] (Theorem 2.1).

Theorem 1. *Suppose that $\mathbf{z}_1, ..., \mathbf{z}_k$ are elements of a convex polyhedron K that is subdivided into r-dimensional simplices $S_1, ..., S_h$ with pairwise disjoint interiors and the set of all of their vertices is equal to $\{\mathbf{z}_1, ..., \mathbf{z}_k\}$. Suppose further that there exists a convex function $f(\mathbf{z}), \mathbf{z} \in K$, continuous on K and linear on any of the simplices $S_1, ..., S_h$ with different normal vectors on different simplieces such that $f_i = f(\mathbf{z}_i), i = 1, ..., k$. Let $B_1, ..., B_h$ be those $(n+1) \times (n+1)$ parts of the matrix of equality constraints of problem (2.3), the upper $n \times (n+1)$ parts of which are the sets of vertices of the simplices $S_1, ..., S_h$, respectively. Then $B_1, ..., B_h$ are the dual feasible bases of problem (2.3) and each of them is dual nondegenerate.*

If the above-mentioned normal vectors are not all different, the assertion that the vertices of any simplex form a dual feasible basis, remains true but these bases are no longer all dual nondegenerate, as it turns out from the proof of the theorem.

Let $S_1,\ ...,\ S_{n!}$ designate the subdividing simplices. Let us rewrite problem (2.5) into more detailed form:

$$
\begin{aligned}
& \min \sum_{k=1}^{N} f(\mathbf{z}_k)\lambda_k \\
\text{subject to} \quad & \\
& x_1 + u_1 = b_1 \\
& \vdots \\
& x_n + u_n = b_n \\
& \sum_{k=1}^{N} \begin{pmatrix} z_{k1} \\ \vdots \\ z_{kn} \end{pmatrix} \lambda_k - \begin{pmatrix} x_1 \\ \vdots \\ 0 \end{pmatrix} - \ldots - \begin{pmatrix} 0 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \\
& \sum_{k=1}^{N} \lambda_k = 1 \\
& \lambda \geq 0\ , \quad \mathbf{x} \geq 0\ , \quad \mathbf{u} \geq 0.
\end{aligned}
\tag{2.6}
$$

It is easy to see that the rank of the matrix of equality constraints in (2.6) is $2n+1$. Let $v_1,\ v_2,\ \ldots\ ,\ v_n,\ y_1,\ y_2,\ \ldots\ ,\ y_n,\ w$ designate the dual variables. The coefficient matrix of problem (2.6) has a special structure illustrated in Table 1.

0	$\cdots$	0	0	$\cdots$	0	$f(\mathbf{z}_1)$	$\cdots$	$f(\mathbf{z}_n)$	
1	$\cdots$	0	1	$\cdots$	0	0	$\cdots$	0	b_1
$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
0	$\cdots$	1	0	$\cdots$	1	0	$\cdots$	0	b_n
-1	$\cdots$	0	0	$\cdots$	0	z_{11}	$\cdots$	z_{N1}	0
$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
0	$\cdots$	-1	0	$\cdots$	0	z_{1n}	$\cdots$	z_{Nn}	0
0	$\cdots$	0	0	$\cdots$	0	1	$\cdots$	1	1
Block 0						Block 1			

Table 1. Coefficient matrix of problem (2.6), together with the objective function coefficients and the right-hand side vector

First, let us introduce the notations:

$$
A = \begin{pmatrix} 1 & \ldots & 0 & 1 & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 1 & 0 & \ldots & 1 \end{pmatrix}, \quad T = \begin{pmatrix} -1 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & -1 & 0 & \ldots & 0 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}
$$

To initiate the dual algorithm we compose a dual feasible basis for problem (2.6). To accomplish this job we pick an arbitrary simplex S_i whose vertices are $\mathbf{z}_{i_1},\ \ldots,\ \mathbf{z}_{i_{n+1}}$ and form the $2n+1$-component vectors

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mathbf{z}_{i_1} \\ 1 \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mathbf{z}_{i_{n+1}} \\ 1 \end{pmatrix}. \tag{2.7}$$

Then we compute the dual variables $\mathbf{y}$ and w by using the equation:

$$\mathbf{z}_{i_k}^T \mathbf{y} + w = f(\mathbf{z}_{i_k}),\ \ k = 1, ..., n+1\ .$$

Next we solve the linear programming problem

$$\begin{aligned} &\min\quad -\mathbf{y}^T T \begin{pmatrix} \mathbf{x} \\ \mathbf{u} \end{pmatrix} \\ &\text{subject to} \\ &\qquad x_1 + u_1 = b_1 \\ &\qquad\qquad \vdots \\ &\qquad x_n + u_n = b_n \\ &\qquad \mathbf{x} \geq 0\ ,\quad \mathbf{u} \geq 0, \end{aligned} \tag{2.8}$$

by a method that produces a primal-dual feasible basis. Let B be this optimal basis and d a dual vector corresponding to B, i.e., any solution of the equation $d^T B = -\mathbf{y}^T T_B$, where T_B is the part of T which correspond to the basis subscripts. Since A has full rank B and d is uniquely determined.

Problem (2.9), however, is equivalent to

$$\begin{aligned} &\min\ \ y_1 x_1 + ... + y_n x_n \\ &\text{subject to} \\ &\qquad x_1 + u_1 = b_1 \\ &\qquad\qquad \vdots \\ &\qquad x_n + u_n = b_n \\ &\qquad \mathbf{x} \geq 0\ ,\quad \mathbf{u} \geq 0 \end{aligned} \tag{2.9}$$

which can be solved easily: if $y_i \leq 0$, we take the column of $\mathbf{x}_i$ otherwise we take the column of $\mathbf{u}_i$ into the basis. We have obtained a dual feasible basis for problem (2.6). It consists of those vectors that trace out B from A and T_B from T in Block 0, furthermore the previously selected vectors (2.7) in Block 1 in Table 1. The dual feasibility is guaranteed by the theorem of Prékopa and Li in [11] (Theorem 2.2).

The next step is to check the primal feasibility of basis. The first n constraint in problem (2.6) ensure that in case of any basis the basic $\mathbf{x}_i,\ \mathbf{u}_j,\ (i \neq j)$ variables are positive, since we have the inequality $\mathbf{b} > 0$. Thus, if the basis is not primal feasible, then only the λ variables can take negative value.

If $\lambda_j < 0$, then the column of $\mathbf{z}_{i_j}$ can be chosen to be the outgoing vector. The incoming vector is either a nonbasic column from Block 0 or a nonbasic column from Block 1 in Table 1. The algorithm can be described in the following way.

Dual algorithm to maximize strongly unimodal functions

Let us introduce the notations:

$$H = \{1, 2, ..., h\}, \quad K = \{1, 2, ..., k\},$$

$$Z_j = \{\mathbf{z}_k \mid \mathbf{z}_k \text{ is a vertex of the simplex } \mathcal{S}_j \ , \ k \in K\}, \quad j \in H,$$

$$L_j = \{k \mid \mathbf{z}_k \in Z_j\}, \quad j \in H.$$

Step 0. Pick arbitrarily a simplex $\mathcal{S}_j$ and let $\mathbf{z}_i, \ i \in \{i_1, ..., i_{n+1}\}$ be the collection of its vertices and

$$I^{(0)} \leftarrow L_j = \{i_1, ..., i_{n+1}\}$$

Go to Step 1.

Step 1. Solve for $\mathbf{y}, \ w$ the system of linear equations:

$$\mathbf{z}_{i_k}^T \mathbf{y} + w = f(\mathbf{z}_{i_k}), \quad k \in I^{(0)},$$

where $\mathbf{y} \in \mathbb{R}^n, \ w \in \mathbb{R}$ Go to Step 2.

Step 2. Compose a dual feasible basis B by including the vectors (2.7), any column of $\mathbf{x}_i$ if $y_i \leq 0$ and any column of $\mathbf{u}_i$ if $y_i > 0$. Go to Step 3.

Step 3. Check the primal feasibility of basis B. If $\lambda \geq 0$, stop, the basis is optimal. If $\lambda \ngeq 0$, then pick $\lambda_q < 0$ arbitrarily and remove $\mathbf{z}_q$ from the basis. Go to Step 4.

Step 4. *Determination of incoming vector.* The following columns may enter:

(1) A nonbasic column from Block 0.

(2) A nonbasic column $\mathbf{z}_j$ from Block 1.

Updating formulas

(1) Given $I^{(k)}$, in order to update a column from Block 0 which traces out the nonbasic column a_p from A, solve the following system of linear equations:

$$\begin{aligned} A_B d_p &= a_p \\ T_B d_p + \sum_{i \in I^{(k)}} \mathbf{z}_i d_i &= t_p \\ \sum_{i \in I^{(k)}} d_i &= 1 \ , \end{aligned} \tag{2.10}$$

where A_B, T_B are the parts of A and T, respectively, corresponding to basis B; d_p is a vector with suitable size and t_p is the p column of the matrix T.

Compute the reduced costs:

$$\bar{c}_p = \sum_{i \in I^{(k)}} f(\mathbf{z}_i) d_i \ .$$

(2) Assume that a nonbasic column $\mathbf{z}_j$ from Block 1, where $|I^{(k)}| < n+1$, $j \neq q$ and $\{j\} \cup I^{(k)} \backslash \{q\}$ is a subset of L_l for some $l \in H$, enters the basis. Let $\hat{I}^{(k)}$ designate the set of all possible j from Block 1 satisfying above requirements. To update the column containing $\mathbf{z}_j$, $j \in \hat{I}^{(k)}$, we solve the system of linear equations:

$$\begin{aligned} A_B r_j &= 0 \\ T_B r_j + \sum_{i \in \hat{I}^{(k)}} \mathbf{z}_i d_i &= \mathbf{z}_j \\ \sum_{i \in \hat{I}^{(k)}} d_i &= 1 \ , \end{aligned} \tag{2.11}$$

where r_j is a vector with suitable size.

Compute the reduced costs:

$$\bar{f}_j = -f(\mathbf{z}_j) \ .$$

<u>*Determination of the vector that enters the basis*</u>

Let $\tilde{d}^{\ T} = (d_p^T, d_{i_1}, ..., d_{i_{n+1}})$ and $\tilde{d}(q)$ be the qth component of $\tilde{d}$ in (2.10). Let $\tilde{r}^{\ T} = (r_j^T, d_{i_1}, ..., d_{i_{n+1}})$ and $\tilde{r}(q)$ be the q component of $\tilde{r}$ in (2.11). Then the incoming vector is determined by taking the minimum of the following minima:

$$min_{\tilde{d}(q)<0} \left\{ \frac{\sum_{i \in I^{(k)}} f(\mathbf{z}_i) d_i}{\tilde{d}(q)} \right\} \ , \tag{2.12}$$

$$min_{\tilde{r}(q)<0} \left\{ \frac{-f(\mathbf{z}_j)}{\tilde{r}(q)} \right\} \ . \tag{2.13}$$

If the minimum is attained in (2.12), let

$$I^{(k+1)} = I^{(k)} \backslash \{q\} \ .$$

Update the basis B by replacing the outgoing vector by the column of a_p in Block 0.

If the minimum is attained in (2.13), then the column of $\mathbf{z}_j$ is the incoming vector. Let

$$I^{(k+1)} = I^{(k)} \cup \{j\} \backslash \{q\} \ .$$

Update the basis B by replacing the outgoing vector by the column of $\mathbf{z}_j$ in Block 1. Go to Step 3.

If no two linear pieces of the function $f(\mathbf{x})$ are on the same hyperplane, then cycling cannot occur, i.e., no simplex that has been used before returns. Otherwise an anti-cycling procedure has to be used: lexicographic dual algorithm (see, e.g., [12]) or Bland's rule [**?**].

We can also find bounds for the optimum value of problem (2.6) by the use of the fast bounding technique by Prékopa and Li [11]. First we construct a dual feasible basis as described before. If $\mathbf{v}$, $\mathbf{y}$, w are the corresponding dual vectors, then $b^T\mathbf{v} + w$ is a lower bound for the optimum value of problem (2.6). In order to find an upper bound we use any method that produces a pair of primal and dual optimal solutions (not necessarily an optimal basis). Having the optimal $(\hat{\mathbf{x}}^T, \hat{\mathbf{u}}^T)$, we arbitrarily pick a simplex $\mathcal{S}_k$ and represent $(\hat{\mathbf{x}}^T, \hat{\mathbf{u}}^T)$ as the convex combination of the vertices of S_k. If all coefficients are nonnegative, then we stop. Otherwise we delete the corresponding vertex from the simplex and update the basis by including the vertex of the neighboring simplex into the basis which is not a vertex of the current simplex. If the representation of the vector $-T(\hat{\mathbf{x}}^T, \hat{\mathbf{u}}^T)$ is

$$-T(\hat{\mathbf{x}}^T, \hat{\mathbf{u}}^T) = \sum_{\mathbf{z}_j \in \mathcal{S}} \mathbf{z}_j \lambda_j,$$

where $\sum_j \lambda_j = 1$, $\lambda_j \geq 0$, then the upper bound is given by

$$-\mathbf{y}^T T(\hat{\mathbf{x}}^T, \hat{\mathbf{u}}^T) + \sum_j f(\mathbf{z}_j)\lambda_j \ .$$

The solution of problem (2.2) can be accomplished in the same way, only trivial modifications are needed. If the minimum of f is taken in the set $\{\mathbf{x} \in \mathbb{Z}^n \mid \mathbf{x} > 0,\ x_1 + ... + x_n \leq b\}$, then we assume that $f(\mathbf{x}) = M$ $(M > 0)$ for every $\mathbf{x}$ that does not belong to this set, where M is large enough (or ∞). In this case, problem (2.2) can be solved by the use of above-mentioned methods.

We also remark that in continuous optimization one of the important properties of convex functions is the coincidence between their local and global minima. A function is $g : \mathbb{Z}^n \to \mathbb{R}$ called integrally convex if and only if its extension $\tilde{g} : \mathbb{R}^n \to \mathbb{R}$ is convex. In this case a global minimum for (continuous) function $\tilde{g}$ is a global minimum for (discrete) function g, and vice versa.

References

[1] O. Barndorff-Nielsen, Unimodality and exponential families, Comm. Statist. 1 (1973), 189–216.

[2] C.I. Fábián, A. Prékopa, O. Ruf-Fiedler, On a dual method for a speacially structured linear programming problem with application to stochastic programming, Optimization Methods and Software 17 (2002), 445–492.

[3] P. Favati, F. Tardella, Convexity in nonlinear integer programming, Ricerca Operativa 53 (1990) 3–44.

[4] M. Fekete, G. Pólya, Überein ein Problem von Laguerre, Rediconti del Circolo Matematico di Palermo 23 (1912), 89–120.

[5] N.L. Johnson, S. Kotz, N. Balakrishnan, Discrete Multivariate Distributions, Wiley, New York, 1997.

[6] K. Murota, Discrete convex analysis, Mathematical Programming 83 (1998), 313–371.

[7] J.G. Pedersen, On strong unimodality of two-dimensional discrete distributions with applications to M-ancillarity, Scand J. Statist. 2 (1975), 99–102.

[8] A. Prékopa, Logarithmic concave functions with applications to stochastic programming, Acta. Sci. Math. 32 (1971), 301–316.

[9] A. Prékopa, On logarithmic concave measures and functions, Acta. Sci. Math. 34 (1973), 335–343.

[10] A. Prékopa, Stochastic Programming, Kluwer Academic Publishers, Dordtecht, Boston, 1995.

[11] A. Prékopa, W. Li, Solution of and bounding in a linearly constrained optimization problem with convex, polyhedral objective function, Mathematical Programming 70 (1995), 1–16.

[12] A. Prékopa, A brief introduction to linear programming, Math. Scientist 21 (1996), 85–111.

[13] E. Subasi, M. Subasi, A. Prékopa, On strong unimodality of multivariate discrete distributions, Discrete Applied Mathematics 157 (2009), 234–246.