APPLICATION OF THE SOLUTION OF THE UNIVARIATE DISCRETE MOMENT PROBLEM FOR THE MULTIVARIATE CASE

GERGELEY MÁDI-NAGY

Department of Differential Equations, Mathematical Institute, Budapest University of Technology and Economics, Müegyetem rakpart 1-3, H-1111 Budapest, Hungary
e-mail: gnagy@math.bme.hu

Communicated by D. Miklós

(Received September 6, 2008; accepted April 16, 2009)

Abstract

The objective of the univariate discrete moment problem (DMP) is to find the minimum and/or maximum of the expected value of a function of a random variable which has a discrete finite support. The probability distribution is unknown, but some of the moments are given. This problem is an ill-conditioned LP, but it can be solved by the dual method developed by Prékopa. The multivariate discrete moment problem (MDMP) is the generalization of the DMP where the objective function is the expected value of a function of a random vector. The MDMP has also been initiated by Prékopa and it can also be considered as an (ill-conditioned) LP. The central results of the MDMP concern the structure of the dual feasible bases and provide us with bounds without any numerical difficulties. Unfortunately, in this case not all the dual feasible bases have been found hence the multivariate counterpart of the dual method of the DMP cannot be developed. However, there exists a method developed by Mádi-Nagy [7] which allows us to get the basis corresponding to the best bound out of the known structures by optimizing independently on each variable. In this paper we present a method using the dual algorithm of the DMP for solving those independent subproblems. The efficiency of this new method is illustrated by numerical examples.

1. Introduction

The theory of the discrete moment problem (DMP) has been discussed in Prékopa [12], [14]. Let $X$ be a random variable with a known, finite support...
\[ Z = \{z_0, z_1, \ldots, z_n\}, \text{ where } z_0 < \cdots < z_n. \] The probability distribution of \( X \) is unknown, but some of the moments of \( X \) are known. Consider a function \( f \) with the domain \( Z \). The objective of the DMP is to yield lower and upper bounds for the expected value of \( f(X) \), using the moment information.

In this paper the power moments are taken into account up to a certain order \( m \). We introduce the following notations:

\[ f_i := f(z_i), \quad p_i := P(X = z_i), \quad i = 0, 1, \ldots, n. \]

\[ \mu_k := E(X^k), \quad k = 0, 1, \ldots, m. \]

Our DMP can be represented as an LP:

\[
\min(\max) \quad E[f(X)] = \{f_0p_0 + f_1p_1 + \cdots + f_np_n\}
\]

subject to

\[
\begin{align*}
p_0 + p_1 + \cdots + p_n &= 1 \quad \text{(1)} \\
z_0p_0 + z_1p_1 + \cdots + z_np_n &= \mu_1 \\
z_0^2p_0 + z_1^2p_1 + \cdots + z_n^2p_n &= \mu_2 \\
& \quad \vdots \\
z_0^mp_0 + z_1^mp_1 + \cdots + z_n^mp_n &= \mu_m \\
p_0 \geq 0, \quad p_1 \geq 0, \quad \ldots, \quad p_n \geq 0,
\end{align*}
\]

where the unknown variables are \( p_i, i = 0, 1, \ldots, n \).

The multivariate discrete moment problem (MDMP) is a generalization of the DMP for random vectors. It has been discussed in papers by Prékopa [13], [15], [16], Mádi-Nagy and Prékopa [9]. Let \( X = (X_1, \ldots, X_s) \) be a random vector with unknown distribution. We assume that the support of \( X_j \) is a known finite set \( Z_j = \{z_{j0}, \ldots, z_{jn_j}\} \) consisting of distinct elements. We define

\[ p_{i_1\ldots i_s} = P(X_1 = z_{i_1}, \ldots, X_s = z_{i_s}), \quad 0 \leq i_j \leq n_j, \quad j = 1, \ldots, s, \]

\[ \mu_{\alpha_1\ldots\alpha_s} = \sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} z_{i_1}^{\alpha_1} \cdots z_{i_s}^{\alpha_s} p_{i_1\ldots i_s}, \]

where \( \alpha_1, \ldots, \alpha_s \) are nonnegative integers. The number \( \mu_{\alpha_1\ldots\alpha_s} \) will be called the \((\alpha_1, \ldots, \alpha_s)\)-order moment of the random vector \((X_1, \ldots, X_s)\), and the sum \( \alpha_1 + \cdots + \alpha_s \) will be called the total order of the moment. Let \( Z = Z_1 \times \cdots \times Z_s \) and \( f(z), z \in Z \) be a function on the domain \( Z \). Let \( f_{i_1\ldots i_s} = \)
In our paper we consider the following MDMP:

$$\min(\max) \ E[f(X)] = \sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} f_{i_1 \ldots i_s} p_{i_1 \ldots i_s}$$

subject to

$$\sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} z_{i_1}^{\alpha_1} \cdots z_{i_s}^{\alpha_s} p_{i_1 \ldots i_s} = \mu_{\alpha_1 \ldots \alpha_s}$$

$$\alpha_j \geq 0, \ j = 1, \ldots, s; \ \alpha_1 + \cdots + \alpha_s \leq m \ \text{or}$$

$$\alpha_j = 0, \ j = 1, \ldots, k - 1, k + 1, \ldots, s,$$

$$m \leq \alpha_k \leq m_k, \ k = 1, \ldots, s;$$

$$p_{i_1 \ldots i_s} \geq 0, \ \text{for all } i_1, \ldots, i_s.$$
The examples are also good illustrations of possible applications of MDMP. Section 5 concludes the paper.

2. Bounds based on dual feasible bases of the univariate and multivariate DMP

The coefficient matrices of the DMPs (1) and (2) are ill-conditioned Vandermonde-type matrices hence the DMP usually cannot be solved by the regular methods and solvers. One way could be the usage of multiple precision arithmetic, but this leads to a very long solution time.

However, in the univariate case, under some assumptions on the function \( f \), all the dual feasible bases of problem (1) can be given. The condition on the function \( f \) is given in terms of divided differences hence we give the following

**Definition 2.1.** Let \( f(z), z \in \{z_0, \ldots, z_n\} \) be a univariate discrete function, where \( z_0, \ldots, z_n \) are distinct real numbers.

\[
[z_i; f] := f(z_i), \quad \text{where} \quad z_i \in \{z_0, \ldots, z_n\}
\]

The \( k \)th order (univariate) divided differences \((k \geq 1)\) are defined recursively:

\[
[z_i, \ldots, z_{i+k}; f] = \frac{[z_{i+1}, \ldots, z_{i+k}; f] - [z_i, \ldots, z_{i+k-1}; f]}{z_{i+k} - z_i},
\]

where \( z_i \in \{z_0, \ldots, z_n\} \).

It is easy to see that if the interval \([z_0, z_n]\) is subset of the domain of the function \( f(z) \) and the function has continuous, positive \( k \)th derivatives in the interior of the interval, then all divided differences of order \( k \) of \( f(z) \), \( z \in Z \) are positive. The theorem concerning the dual feasible bases is the following.

**Theorem 2.1** (Prékopa [14], Th. 5.7.1.). Suppose that all \( m+1 \)st divided differences of the function \( f(z), z \in \{z_0, z_1, \ldots, z_n\} \) \((z_0 < \cdots < z_n)\) are positive. Then in problem (1), all bases are dual-nondegenerate and the dual feasible bases have the following structures, presented in terms of the subscripts of the basis vectors:

\[
m + 1 \text{ even} \quad m + 1 \text{ odd}
\]

\[
\begin{align*}
\min & \quad \{j, j+1, \ldots, k, k+1\} & \{0, j, j+1, \ldots, k, k+1\} \\
\max & \quad \{0, j, j+1, \ldots, k, k+1, n\} & \{j, j+1, \ldots, k, k+1, n\}
\end{align*}
\]

I.e., the set of subscripts of the basis variables consists of pairs of consecutive elements completed by 0 and/or \( n \) depending on its parity and its \( (\min \text{ or } \max) \) type.
Based on the above theorem a numerically stable dual method was developed by Prékopa [12]. This method, fitted to a certain subproblem of the multivariate DMP, is described in Section 3. An illustrative step-by-step example of this method can be found in Prékopa [14], pp. 167–168.

In case of the multivariate DMP not all the dual feasible bases are known hence we cannot construct a robust dual simplex method to solve it. However, some dual feasible basis structures can be given and by the aid of them bounds can be derived for (2). Furthermore, if the number of the known dual feasible bases is large enough, then the best corresponding bounds are close to the optimal values (min and max) of the MDMP (2).

We will use the notations of the compact matrix form of problem (2) (compatible with the notation of Mádi-Nagy and Prékopa [9] and Mádi-Nagy [7]):

\[
\begin{align*}
\min \text{ (max) } & \quad f^T p \\
\text{subject to } & \quad \tilde{A} p = \tilde{b} \\
& \quad p \geq 0.
\end{align*}
\]

Let \( V_{\min} \) (\( V_{\max} \)) designate the minimum (maximum) value in problem (4). Let further \( B_1 \) (\( B_2 \)) designate a dual feasible basis for the minimization (maximization) problem. Then, by linear programming theory, we know that

\[
\begin{align*}
f_{B_1}^T p_{B_1} \leq V_{\min} \leq E[f(X_1, \ldots, X_s)] = f^T p \leq V_{\max} \leq f_{B_2}^T p_{B_2}.
\end{align*}
\]

I.e., the objective function value of a dual feasible basic solution of problem (4) always gives bound for \( E[f(X_1, \ldots, X_s)] \). The following theorems give bounds for the objective function by the aid of this relation. First, they define dual feasible bases and then, based on (5), they also give bounds.

In the following definitions of dual feasible bases we always use the set of subscripts

\[
I = I_0 \cup \left( \bigcup_{j=1}^{s} I_j \right),
\]

where

\[
I_0 = \{ (i_1, \ldots, i_s) | 0 \leq i_j \leq m - 1, \text{ integers, } j = 1, \ldots, s, \ i_1 + \cdots + i_s \leq m \}
\]

and

\[
I_j = \{ (i_1, \ldots, i_s) | i_j \in K_j, \ i_l = 0, \ l \neq j \}
\]

\[
K_j = \{ k_j^{(1)}, \ldots, k_j^{(|K_j|)} \} \subset \{m, m + 1, \ldots, n_j\}, \ j = 1, \ldots, s.
\]
We shall consider four different structures for $K_j$:

\[(9)\]

\[
\begin{align*}
|K_j| \text{ even} & \quad \min u^{(j)}, u^{(j)} + 1, \ldots, v^{(j)}, v^{(j)} + 1 \\
|K_j| \text{ odd} & \quad \max m, u^{(j)}, u^{(j)} + 1, \ldots, v^{(j)}, v^{(j)} + 1
\end{align*}
\]

\[
\begin{align*}
\text{min} & \quad u^{(j)}, u^{(j)} + 1, \ldots, v^{(j)}, v^{(j)} + 1 \\
\text{max} & \quad m, u^{(j)}, u^{(j)} + 1, \ldots, v^{(j)}, v^{(j)} + 1, n_j \\
\text{min} & \quad u^{(j)}, u^{(j)} + 1, \ldots, v^{(j)}, v^{(j)} + 1 \\
\text{max} & \quad m, u^{(j)}, u^{(j)} + 1, \ldots, v^{(j)}, v^{(j)} + 1, n_j.
\end{align*}
\]

We also need the following

**Definition 2.2.** Let $f(z), z \in Z = Z_1 \times \cdots \times Z_s$ be a multivariate discrete function and take the subset

\[(10)\]

\[
Z_{I_1, \ldots, I_s} = \{z_{1i}, i \in I_1\} \times \cdots \times \{z_{si}, i \in I_s\}
\]

where $|I_j| = k_j + 1, j = 1, \ldots, s$. Then we can define the $(k_1, \ldots, k_s)$-order \textit{multivariate} divided difference of $f$ on the set \((10)\) in an iterative way. First we take the $k_1$th divided difference with respect to the first variable, then the $k_2$th divided difference with respect to the second variable etc. These operations can be executed in any order even in a mixed manner, the result being always the same. Let

\[(11)\]

\[
\begin{align*}
[z_{1i}, i \in I_1; \ldots; z_{si}, i \in I_s; f]
\end{align*}
\]

designate the $(k_1, \ldots, k_s)$-order divided difference. The sum $k_1 + \cdots + k_s$ is called the total order of the divided difference.

In order to make the definition easier to understand we present the following

**Example 2.1.** We would like to calculate the $(1, 1)$-order divided difference of the bivariate function $f(z_1, z_2)$:

\[
\begin{align*}
[z_{10}, z_{11}; z_{20}, z_{21}; f].
\end{align*}
\]

First, we look at $f$ as a function of the first variable $z_1$ (assumed to be fixed) and take the first divided difference respect to $z_{10}, z_{11}$, i.e.

\[
\begin{align*}
[z_{10}, z_{11}; f(z_{11}, z_2)] = \frac{f(z_{11}, z_2) - f(z_{10}, z_2)}{z_{11} - z_{10}}.
\end{align*}
\]

Then we look at the result as a function of $z_2$ and take the first divided difference respect to $z_{20}, z_{21}$, i.e.

\[
\begin{align*}
[z_{20}, z_{21}; f(z_{11}, z_2) - f(z_{10}, z_2)] = \frac{f(z_{11}, z_2) - f(z_{10}, z_2)}{z_{11} - z_{10}} \quad \frac{f(z_{11}, z_{20}) - f(z_{10}, z_{20})}{z_{21} - z_{20}}.
\end{align*}
\]
UNIVARIATE DISCRETE MOMENT PROBLEM

Thus

\[
[z_{10}, z_{11}; z_{20}, z_{21}; f] = \frac{f(z_{21}, z_{20})-f(z_{11}, z_{21})}{z_{21}-z_{10}} - \frac{f(z_{21}, z_{10})-f(z_{11}, z_{10})}{z_{21}-z_{10}}.
\]

If \( f(z) \), \( z \in Z \) is derived from a function \( \tilde{f}(z) \) defined in \( Z = [z_{10}, z_{1} \times \cdots \times [z_{s0}, z_{s}] \) by taking \( f(z) = \tilde{f}(z), z \in Z \) and \( \tilde{f}(z) \) has continuous, nonnegative derivatives of order \( (k_1, \ldots, k_s) \) in the interior of \( Z \), then all divided differences of \( f(z), z \in Z \) of order \( (k_1, \ldots, k_s) \) are nonnegative. For further results in this respect see Popoviciu [11].

Then we consider the following

**Assumption 2.1.** The function \( f(z), z \in Z \) has nonnegative divided differences of total order \( m + 1 \), and in addition, in each variable \( z_j \) it has positive divided differences of order \( m_j + 1 := m + |K_j| \).

We remark that for the results of this section it is enough to assume the nonnegativity of the univariate moments. Positivity is needed in the following section of the new algorithm.

**Theorem 2.2** (Mádi-Nagy and Prékopa [9]). Let \( z_{j0} < z_{j1} < \cdots < z_{jn_j}, j = 1, \ldots, s \). Suppose that function \( f \) fulfils Assumption 2.1, where the set \( K_j \) has one of the min structures in (9).

Under these conditions the set of columns \( \hat{B} \) of \( \hat{A} \) in problem (4), with the subscript set \( I \), is a dual feasible basis in the minimization problem (4), and

\[
E[f(X_1, \ldots, X_s)] \geq f^T \hat{p} \hat{B}.
\]

If \( \hat{B} \) is also a primal feasible basis in problem (4), then the inequality (12) is sharp.

**Theorem 2.3** (Mádi-Nagy and Prékopa [9]). Let \( z_{j0} > z_{j1} > \cdots > z_{jn_j}, j = 1, \ldots, s \). Suppose that function \( f \) fulfils Assumption 2.1, where \( K_j \) has one of the structures in (9) that we specify below.

(a) If \( m + 1 \) is even, \( |K_j| \) is even and \( K_j \) has the max structure in (9) or \( m + 1 \) is even, \( |K_j| \) is odd and \( K_j \) has the min structure in (9), then the set of columns \( \hat{B} \) in \( \hat{A} \), corresponding to the subscripts \( I \), is a dual feasible basis in the minimization problem (4). We also have the inequality

\[
E[f(X_1, \ldots, X_s)] \geq f^T \hat{p} \hat{B}.
\]

(b) If \( m + 1 \) is odd, \( |K_j| \) is even and \( K_j \) has the max structure in (9) or \( m + 1 \) is odd, \( |K_j| \) is odd and \( K_j \) has the min structure in (9), then
the basis $\hat{B}$ is dual feasible in the maximization problem (4). We also have the inequality

$$E[ f(X_1, \ldots, X_s)] \leq f^T_B p_B.$$  

The two above theorems yield dual feasible basis structures by the aid of the subscript set $I$ defined in (6), ordering the elements of $Z_j$’s increasingly or decreasingly. In the bivariate case ($s = 2$), (still at Assumption 2.1) we can give much more dual feasible bases corresponding to $I$, by suitable (not necessary increasing or decreasing) ordering of the variables. Detailed discussion with illustrative figures and examples can be found in Mádi-Nagy and Prékopa [9]. The multivariate generalization of these algorithms can be found in Mádi-Nagy [8] where the MDMP is slightly different to (2) but the algorithm remains nearly the same. Hence, the further results of our paper can be applied for the MDMP of Mádi-Nagy [8], as well.

3. Application of the univariate method in the solution of the MDMP

In the theorems of the previous section we dealt with bases corresponding to the subscript set $I$ (6). Let us call them $Z_I$-type bases. Our aim is to find the dual feasible basis among them that gives the maximum (minimum) objective value function in case of the minimum (maximum) problem of (2), i.e., to find the basis that gives the closest bound among them. The diversity of $Z_I$-type bases is given by the order of $Z_j$’s and the choices of the subscript sets $K_j$’s.

If the orders of $Z_j$’s are given, then by the method of Mádi-Nagy [7] the best $K_j$’s, in sense of the objective function value, can be found independently. The method of our paper, furthermore, can find those best $K_j$’s by a dual method similar to the univariate method of Prékopa [12]. This has two advantages. On one hand, instead of considering all possible $K_j$’s and the corresponding objective function values, at every step we get a $K_j$ which gives a better bound until the $K_j$ which corresponds to the best bound. On the other hand, at every step the method yields the next $K_j$ without calculating the inverse of a Vandermonde-type basis. This means much less calculation as well as more numerical stability. In the following the new method is described. We just sketch the part of the method which is identical of the method of Mádi-Nagy [7] and go into details only at the novelties.

All MDMPs with Assumption 2.1 can be converted into an equivalent problem, such that Assumption 2.1 remains valid and

$$z_{j0} = 0, \quad j = 1, \ldots, s$$
and

\[ f(z_{10}, \ldots, z_{s0}) = 0. \]

Consider the following subscript sets:

\[ I^\text{int}_0 = \{ (i_1, \ldots, i_s) \mid 1 \leq i_j \leq m - 1, \text{ integer}, \]
\[ j = 1, \ldots, s, \ i_1 + \cdots + i_s \leq m \}, \]
\[ I^\text{axes}_j = \{ (i_1, \ldots, i_s) \mid 1 \leq i_j \leq n_j, \text{ integer}; i_l = 0 \text{ for } l \neq j \}. \]

If we reorder the columns and rows of the constraint matrix of the converted problem according to the subscript sets above, we get a more perspicuous structure:

\[ f^T : \begin{array}{cccc}
0 & z_{1}^{\text{xaxes}} & \ldots & \hat{b} \\
0 & f_{1}^{T \text{xaxes}} & \ldots & f_{1}^{T \text{int}} \\
0 & f_{2}^{T \text{xaxes}} & \ldots & f_{2}^{T \text{int}} \\
& \vdots & \ddots & \vdots \\
0 & z_{s}^{\text{xaxes}} & \ldots & z_{s}^{\text{int}} \\
0 & f_{s}^{T \text{xaxes}} & \ldots & f_{s}^{T \text{int}} \\
0 & f_{s+1}^{T} & \ldots & f_{s+1}^{T} \\
0 & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0
\end{array} \]

\[ \hat{A} : \begin{array}{cccc}
1 & 1 & \cdots & 1 \\
0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0
\end{array} \]

At problem (18) we have introduced some new notations, which help us in the following arguments.
The subscripts denote the columns of the matrix, while the superscripts refer to the rows.

- $p$ denotes an appropriate basic solution, while the $p_B$ vector consists of the components of the basic variables. From the structure of the $Z_I$ bases follows that there are no basic variables in the last block of $p^T$ hence these components have the value zero.

- Between the rows $p^T_B$ and $p^T$ we refer by equality signs to the fact, that all variables of $p_{0...0}$ and $p_{0...0}^{I_0}$ are also basic variables for each $Z_I$-type basic solution.

First, let us consider the rows of the mixed moments. Since the variables of the last block are zero and the coefficients of the variables of $Z_{I_0}$ are zero except $Z_{I_0}^{I_0}$ and the matrix $A_{I_0}^{I_0}$ is a non-singular square matrix:

\begin{equation}
\label{eq:21}
p_{I_0}^{I_0} = \left( A_{I_0}^{I_0} \right)^{-1} I_{I_0}^{I_0}.
\end{equation}

Let $\hat{b}_I = \hat{b} - A_{I_0}^{I_0} p_{I_0}^{I_0}$. Then the problem is broken into the following type of smaller subproblems:

\begin{equation}
\label{eq:22}
\begin{align*}
\text{max} & \quad (\text{min}) \quad f_{I_j}^{\text{axes}} p_{I_j}^{\text{axes}} \\
\text{subject to} & \\
A_{I_j}^{\text{axes}} p_{I_j}^{\text{axes}} & = \hat{b}_I^{I_j} \\
p_{I_j}^{\text{axes}} & \geq 0
\end{align*}
\end{equation}

$j = 1, \ldots, s$. The last constraint above means that the subscript set of the basic variables of $p_{I_j}^{\text{axes}}$ is the union of the part of $I_0$, which contains the related axis and the set $I_j$ that is characterized by $K_j$. In our new method we focus on subproblems (20). Those give the subscript sets $K_j$'s corresponding to the $Z_I$-type basis yielding the best bound.

First we prove

**Theorem 3.1.** The corresponding parts of $Z_I$-type bases, $B_{I_j}^{I_{axes}}$'s, are dual feasible in problem

\begin{equation}
\label{eq:23}
\begin{align*}
\text{min} & \quad (\text{max}) \quad f_{I_j}^{\text{axes}} p_{I_j}^{\text{axes}} \\
\text{subject to} & \\
A_{I_j}^{\text{axes}} p_{I_j}^{\text{axes}} & = \hat{b}_I^{I_j} \\
p_{I_j}^{\text{axes}} & \geq 0
\end{align*}
\end{equation}
Proof. Let us consider the following problem:

\[
\min (\max) \quad (f_0, f_{I_j}^T) \cdot \begin{pmatrix} p_0 \\ p_{I_j} \end{pmatrix}
\]

subject to

\[
\begin{pmatrix} 1 & 1^T \\ 0 & \tilde{A}_{I_j}^{axos} \end{pmatrix} \cdot \begin{pmatrix} p_0 \\ p_{I_j} \end{pmatrix} = \begin{pmatrix} b_0 \\ b_{I_j} \end{pmatrix},
\]

where \( f_0 = 0 \). The coefficient matrix is, in fact, a Vandermonde matrix, i.e.,

\[
\begin{pmatrix} 1 & 1^T \\ 0 & \tilde{A}_{I_j}^{axos} \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ z_{j_0} & z_{j_1} & \cdots & z_{j_{n_j}} \\ \vdots & \vdots & \ddots & \vdots \\ z_{j_0}^{m_j} & z_{j_1}^{m_j} & \cdots & z_{j_{n_j}}^{m_j} \end{pmatrix},
\]

where \( z_{j_0} = 0 \). The function \( f(z) \) fulfils Assumption 2.1 hence the \( m_j + 1 \)st divided differences of \( f(0, \ldots, 0, z_{j_0}, \ldots, 0) \) in \( z_{j_0} \in Z_j \) are positive.

If we look at the function \( f(0, \ldots, 0, z_{j_0}, \ldots, 0) \) as a univariate function of \( z_{j_0} \in Z_j \), then problem (22) is a univariate DMP and from Theorem 2.1 follows that the corresponding part of any \( Z_{I_j} \)-type basis is dual feasible. We know that the first column (the corresponding part of the column of the variable \( (z_{j_0}, \ldots, z_{j_0}) = (0, \ldots, 0) \)) is always in the corresponding part of a \( Z_{I_j} \)-type basis. Hence, the basis \( B \) can be written in the following form:

\[
B = \begin{pmatrix} 1 & 1^T \\ 0 & \tilde{B}_{I_j}^{axos} \end{pmatrix},
\]

where \( \tilde{B}_{I_j}^{axos} \) is the corresponding part of the same \( Z_{I_j} \)-type basis regarding problem (21).

Finally, we prove that the dual feasibility of the basis \( \tilde{B}_{I_j}^{axos} \) in problem (21) follows from the dual feasibility of the basis \( B \) in problem (22). In case of \( \min (\max) \) problem (22) the dual feasibility of \( B \) means that

\[
(0, f_{I_j}^T) \cdot B^{-1} \cdot \begin{pmatrix} z_{j_0} \\ \vdots \\ z_{j_{n_j}}^{m_j} \end{pmatrix} - f(0, \ldots, 0, z_{j_0}, \ldots, 0) \leq (\geq) 0.
\]
for all \( i = 0, \ldots, n_j \). It is easy to see that

\[
B^{-1} = \begin{pmatrix} 1 & -1^T (\tilde{B}_{f_{I}^{\text{axes}}}^{-1}) \end{pmatrix}.
\]

From this the left-hand side of (23) equals:

\[
(0, f_{I_{j}^{\text{axes}}}) \cdot \begin{pmatrix} 1 & -1^T (\tilde{B}_{f_{I}^{\text{axes}}}^{-1}) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ z_{j1} \\ \vdots \\ z_{jm_j} \\ z_{j1} \end{pmatrix} - f(0, \ldots, 0, z_{j1}, 0, \ldots 0) = f_{I_{j}^{\text{axes}}} \cdot (\tilde{B}_{f_{I}^{\text{axes}}}^{-1}) \cdot \begin{pmatrix} z_{j1} \\ \vdots \\ z_{jm_j} \\ z_{j1} \end{pmatrix} - f(0, \ldots, 0, z_{j1}, 0, \ldots 0).
\]

This means that basis \( \tilde{B}_{f_{I}^{\text{axes}}} \) is dual feasible in problem (21) \( \square \)

Remark 3.1. As regards the choice of \( K_{j} \), it is easy to see, that the corresponding parts of \( Z_{I} \)-type bases are the only dual feasible bases of problem (22) within the case where \( z_{j0}, \ldots, z_{j(m-1)} \) are basic variables. Considering (24) the same is true for problem (21) within the case where \( z_{j1}, \ldots, z_{j(m-1)} \) are basic variables.

From the construction of \( Z_{I} \) follows that after the conversion corresponding to (15) and (16) either all \( z_{ji} \) are positive or all \( z_{ji} \) are negative for \( 1 \leq i \leq n_j \).

If we substitute \( x_{i} = |z_{ji}| p_{0}\cdots 0_{i}0\cdots 0 \) into (21), then we have the following equivalent problem:

\[
\min \ (\max) \ \frac{f_{0\ldots 010\ldots 0}}{|z_{j1}|} x_{1} + \frac{f_{0\ldots 020\ldots 0}}{|z_{j2}|} x_{2} + \cdots + \frac{f_{0\ldots 0n_{j}0\ldots 0}}{|z_{jn_{j}}|} x_{n_{j}}
\]

subject to

\[
\begin{align*}
& x_{1} + x_{2} + \cdots + x_{n_{j}} = b_{f_{I}^{\text{axes}}} & \cdot \arg (z_{j1}) \\
& z_{j1} x_{1} + z_{j2} x_{2} + \cdots + z_{jn_{j}} x_{n_{j}} = b_{f_{I}^{\text{axes}}} & \cdot \arg (z_{j1}) \\
& \vdots \\
& z_{jm_j}^{-1} x_{1} + z_{jm_j}^{-1} x_{2} + \cdots + z_{jm_j}^{-1} x_{n_{j}} = b_{f_{I}^{\text{axes}}} & \cdot \arg (z_{j1}) \\
& x \geq 0
\end{align*}
\]
Considering the results above we can elaborate the following method in order to solve (20). We look for the best corresponding $Z_I$-type basis of the equivalent problem (25) similarly as in the univariate dual method of Prékopa. The new method is based on the result of Theorem 3.1 and Remark 3.1 and the coefficient matrix of (25).

Partial dual method for finding the solution of the subproblem (20), i.e., for finding the best $K_j$

**Step 1:** Pick the corresponding part of any $Z_I$-type basis in agreement with Theorem 2.2 (2.3) or Mádi-Nagy and Prékopa [9]. Let $I_B = \{1, \ldots, m - 1, i_0, i_1, \ldots, i_{m_j - m}\}$ designate the set of subscripts of the basis vectors, where $m \leq i_0, i_1, \ldots, i_{m_j - m} \leq n_j$. Let $K = 0, \ldots, m_j - m$. (The set $K$ is different to the set $K_j$.)

**Step 2:** Determination of the outgoing vector: Take any element $i_k$, $k \in K$. It can be derived (based on Prékopa, [12]) that the sign of the value of the following form equals the sign of the value of the basic variable (i.e., $x_{i_k}$ as well as $p_{i_k}$).

$$(-1)^{m_j-(q_j+k+1)} \times \left[ b_{i_k}^{j_{\text{axes}}} - \left( \sum_{J \in J_{j_{\text{axes}} \backslash \{i_k\}}} z_{ji_j} b_{m_j-1}^{j_{\text{axes}}} + \cdots \right) + (-1)^{m_j-1} \left( \prod_{J \in J_{j_{\text{axes}} \backslash \{i_k\}}} z_{i_j} b_{1}^{j_{\text{axes}}} \right) \right],$$

where $q_j$ is the parameter of the Min (Max) Algorithm. Note that in case of Theorem 2.2 $q_j = m - 1$, in case of Theorem 2.3 $q_j = -1$. Hence, if the value of (26)

- is negative, then the $i_k^\text{th}$ vector of the basis can be the outgoing vector
- is nonnegative, then seek another basis subscript.
- If the value of (26) is positive for all basis subscripts $i_k, k \in K$, then go to Step 4.

**Step 3:** If the outgoing vector is identified, then we can choose at most one incoming vector which restores the $Z_I$-type structure of the basis. If we found the incoming vector, then consider the new $Z_I$-type basis and $K := 0, \ldots, m_j - m$, else $K := K \backslash \{k\}$. Go to Step 2.
Step 4: Stop, we have found the solution of (20), i.e., the corresponding part of $Z_I$-type solution that gives the best bound.

Finally, if we solved problems (20) for $j = 1, \ldots, s$, then

$$p_0^{opt} = (\widehat{b}_1)^0 - (1, \ldots, 1)^T p_{(I_1^{axes}, \ldots, I_s^{axes})}^{opt}.$$ 

By the aid of this method the closest $Z_I$-type bounds can be found in a much shorter way than by calculating the objective function value for all possible $Z_I$-type bases.

The advantage of this new method is that we find the best basis of the subproblem (20) through bases having greater (smaller) objective function values in case of min (max) problem in each step, i.e., we do not have to examine all of the possible bases. In addition we do not have to calculate either the objective function value or the inverse of the basis matrix, just the value of (26).

Finally we summarize the steps of the new method:

(i) Convert the problem into the equivalent form, which satisfies (15) and (16). Then we start to solve this equivalent problem in the following way.

(ii) From equation (19) we know that $p_{I_0^{int}} = (\widehat{A}_{I_0^{int}}^{-1})^T \mu_{I_0^{int}}$.

(iii) Let $\widehat{b}_1 = \widehat{b} - \widehat{A}_{I_0^{int}} p_{I_0^{int}}$. Then we can write up problems (20).

(iv) (The novelty is here!) We solve problems (20) for $j = 1, \ldots, s$, by the application of the partial dual method, described above. Let the optimal solutions be signed by $p_{I_j^{axes}}^{opt}, j = 1, \ldots, s$.

(v) Let $p_0^{opt} = (\widehat{b}_1)^0 - (1, \ldots, 1)^T p_{(I_1^{axes}, \ldots, I_s^{axes})}^{opt}$.

(vi) The other variables are zero, thus we have obtained the closest $Z_I$-type basic solution of the converted problem:

$$(p^{opt})^T = (p_0, (p_{(I_1^{axes}, \ldots, I_s^{axes})}^{opt})^T, 0).$$

(vii) This solution is also the closest $Z_I$-type basic solution of the original problem (2), just the related objective function value will be greater by $f(z_{10}, \ldots, z_{s0})$. 
4. Numerical examples

In this section we present the efficiency of the above method. Three typical applications of the MDMP are considered. First, bounds of the probability of the union of events are given. Then expected utilities are bounded. Finally the values of bivariate moment generating functions are approximated.

These problems cannot be solved by CPLEX and most of them cannot be solved by any general purpose solver, either. The best lower and upper bounds of Mádi-Nagy and Prékopa [9] are given. In order to find the best bounds (best $K_j$s) the method of Mádi-Nagy [7] (which is much faster than the original method) as well as our new method are used. The related CPU times (denoted by $CPU_n$ and $CPU_u$, respectively) are also shown. The algorithms are implemented in Wolfram’s Mathematica 5.1 (www.wolfram.com). Comparing the running times we can see how much faster and more effective the method of our paper is. This also means that by the aid of the new method greater sized problems can be solved.

Example 4.1 (Bounding probability of the union of events). We consider 40 events subdivided into two 20-element groups; $X_j$ equals the number of events that occur in the $j$th group, $j = 1, 2$, $Z_1 \times Z_2 = \{0, \ldots, 20\} \times \{0, \ldots, 20\}$.

We want to find bounds for the probability that at least one out of the 40 events occurs, i.e.

$$P(X_1 + X_2 \geq 1) = E\left[f(X_1, X_2)\right],$$

where

$$f(z_1, z_2) = \begin{cases} 0, & \text{if } (z_1, z_2) = (0, 0) \\ 1, & \text{otherwise.} \end{cases} \quad (27)$$

Prékopa [15] has shown, that if $m + 1$ is even (odd), then all divided differences (27) of total order $m + 1$ are nonpositive (nonnegative).

In case of event sequences usually the so-called cross binomial moments are given. The cross binomial moment of order $(\alpha_1, \ldots, \alpha_s)$ ($\alpha_1, \ldots, \alpha_s$ are nonnegative integers) is defined as:

$$S_{\alpha_1\ldots\alpha_s} = E\left[\frac{X_1^{\alpha_1}}{\alpha_1} \cdots \frac{X_s^{\alpha_s}}{\alpha_s}\right] = \sum_{\substack{1 \leq t_1 < \cdots < t_s \leq n_i \\ j=1,\ldots,s}} P\left[A_{1t_11} \cap \cdots \cap A_{1t_1\alpha_1} \cap \cdots \cap A_{si\alpha_1} \cap \cdots \cap A_{si\alpha_s}\right].$$
Fortunately, if the (cross) binomial moments are given up to a certain order then they can be transformed into power moments up to the same order by the aid of Stirling numbers of the second kind. For details see p. 154 in Prékopa [14]. Hence, in order to give bounds for the probability of the union, the MDMP with function (27) has to be solved.

In our example suppose that we know the following cross binomial moments:

<table>
<thead>
<tr>
<th>1st group</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>742</td>
<td>9961</td>
<td>7021</td>
<td>125489</td>
<td>139384</td>
<td>79478</td>
<td>319816</td>
<td>113127</td>
</tr>
<tr>
<td>1</td>
<td>49</td>
<td>3451</td>
<td>9247</td>
<td>97699</td>
<td>873523</td>
<td>3886421</td>
<td>150</td>
<td>25</td>
<td>75</td>
</tr>
<tr>
<td>2</td>
<td>1524</td>
<td>136549</td>
<td>151893</td>
<td>1281063</td>
<td>11444741</td>
<td>25</td>
<td>75</td>
<td>25</td>
<td>150</td>
</tr>
<tr>
<td>3</td>
<td>40691</td>
<td>283052</td>
<td>1878331</td>
<td>7888804</td>
<td>75</td>
<td>150</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>60888</td>
<td>275947</td>
<td>435869</td>
<td>75</td>
<td>150</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>8004</td>
<td>181702</td>
<td>75</td>
<td>150</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>232106</td>
<td>75</td>
<td>150</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>215523</td>
<td>50</td>
<td>150</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>244503</td>
<td>50</td>
<td>150</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We remark that they are the moments of an event system generated randomly by the method of the numerical examples in Kuai et al. [6]. In that system the probability of the union was 0.7.

Several cases are considered depending on the maximum of the total order of the moments \( m \) and on the maximum of the order of the univariate moments \( m_1, m_2 \) taken into account. The results are the following:

<table>
<thead>
<tr>
<th>( m )</th>
<th>( m_1 )</th>
<th>( m_2 )</th>
<th>Lower</th>
<th>( CPU_n )</th>
<th>( CPU_u )</th>
<th>Upper</th>
<th>( CPU_n )</th>
<th>( CPU_u )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>6</td>
<td>6</td>
<td>0.639566</td>
<td>3.171</td>
<td>0.25</td>
<td>1</td>
<td>3.09</td>
<td>0.23</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>6</td>
<td>0.699931</td>
<td>2.45</td>
<td>0.90</td>
<td>0.763354</td>
<td>2.29</td>
<td>0.90</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>8</td>
<td>0.699999</td>
<td>29.12</td>
<td>1.14</td>
<td>1</td>
<td>29.40</td>
<td>1.34</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>8</td>
<td>0.699999</td>
<td>23.06</td>
<td>4.28</td>
<td>0.755894</td>
<td>22.53</td>
<td>4.46</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>8</td>
<td>0.699999</td>
<td>19.31</td>
<td>10.45</td>
<td>0.707061</td>
<td>21.32</td>
<td>10.54</td>
</tr>
</tbody>
</table>

We can see, if the cardinalities of the sets \( K_j \) (i.e., \( m_j - m + 1 \)) are greater then our new algorithm is more effective. E.g. in case of \( m = 2, m_1 = m_2 = 8 \) the new method is about 20 times faster. This is because in case of large \( |K_j| \) the number of possible bases in subproblems (20) is also
great and the new method solves these problems more effectively. We can also see, that in case of the smallest $|K_j|$s the new method remains at least twice as fast.

**Example 4.2 (Bounding expected multiattribute utility).** Let

\[ f(z_1, z_2) = \log \left( (e^{0.75z_1+2} - 1)(e^{1.25z_2+3} - 1) \right). \]

This function is a member of the utility function class of the paper by Prékopa and Mádi-Nagy [18]. It has the property that its odd (even) order derivatives are nonnegative (nonpositive) for all $z_1 \geq 0$, $z_2 \geq 0$.

Let $Z_1 = Z_2 = \{0, \ldots, 200\}$, $m = 6$, $m_1 = m_2 = 4$. First, we define the moments by the use of discrete uniform distribution on $Z$. In the second case the moments are generated by the use of the following random vector:

\[ \left( \min (X + Y_1, 200), \min (X + Y_2, 200) \right), \]

where $X, Y_1, Y_2$ are random variables having independent Poisson distributions with parameters 30, 40, 50, respectively.

The results, corresponding to the above moments and function (28), are shown below.

<table>
<thead>
<tr>
<th>Moments</th>
<th>Lower</th>
<th>CPU$_n$</th>
<th>CPU$_u$</th>
<th>Upper</th>
<th>CPU$_n$</th>
<th>CPU$_u$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
<td>204.9826</td>
<td>28.73</td>
<td>5.61</td>
<td>205.0000</td>
<td>28.72</td>
<td>5.54</td>
</tr>
<tr>
<td>Poisson</td>
<td>157.4995</td>
<td>23.78</td>
<td>4.61</td>
<td>157.5000</td>
<td>23.61</td>
<td>4.56</td>
</tr>
</tbody>
</table>

The new method turned out to be about five times faster.

**Example 4.3 (Approximating the value of a multivariate moment generating function).** We consider Example 5.1 of Mádi-Nagy and Prékopa [10] where the objective is to give lower and upper bounds for the bivariate moment generating function

\[ M(t_1, t_2) = E[e^{t_1 X_1 + t_2 X_2}], \]

where $t_1 = 0.04$ and $t_2 = 0.05$. 
Let \( m = 4 \), \( m_1 = m_2 = 6 \), and assume that the random vector \( X \) has the following moments:

\[
\begin{array}{cccccccc}
\mu_{ij} & 0 & 1 & 2 & 3 & 4 & 5 & 6 \\
0 & 1 & 7 & 203/3 & 735 & 127687/15 & 102655 & 3818459/3 \\
1 & 7 & 49 & 1421/3 & 5145 & & & \\
2 & 203/3 & 1421/3 & 41209/9 & & & & \\
3 & 735 & 5145 & & & & & \\
4 & 127687/15 & & & & & & \\
5 & 102655 & & & & & & \\
6 & 3818459/3 & & & & & & \\
\end{array}
\]

These values are generated by the use of the discrete uniform distribution on \( \{0, \ldots, 14\} \times \{0, \ldots, 14\} \). Let the support \( Z = Z_1 \times Z_2 \), where \( Z_1 = Z_2 = \{0, \Delta z, 2\Delta z, \ldots, 14\} \). The results, depending on the value of \( \Delta z \), are given in the following tableau:

\[
\begin{array}{cccccc}
\Delta z & \text{Lower} & CPU_n & CPU_u & \text{Upper} & CPU_n & CPU_u \\
1 & 1.95009544 & 1.14 & 0.53 & 1.95107086 & 1.09 & 0.5 \\
0.5 & 1.95003434 & 0.99 & 0.78 & 1.95114443 & 1.01 & 0.73 \\
0.2 & 1.94999630 & 3.02 & 1.80 & 1.95119045 & 3.02 & 1.76 \\
0.1 & 1.94998339 & 6.91 & 3.59 & 1.95120611 & 6.92 & 3.52 \\
0.05 & 1.94997689 & 15.97 & 7.54 & 1.95121400 & 16.33 & 7.51 \\
0.01 & 1.94997167 & 148.16 & 59.65 & 1.95122034 & 148.69 & 63.13 \\
\end{array}
\]

The last row in the tableau illustrates that in case of large problems we can spare a lot of time even in case of small cardinality of the sets \( K_j \).

5. Conclusion

In Theorem 3.1 dual feasible bases have been given for the problem (21), which arises as a subproblem of the method of Mádi-Nagy [7]. Based on this we have developed a numerically stable partial dual method for the solution of problem (20). These results, on one hand, illustrate the connection between the univariate and multivariate DMPs. On the other hand, as the numerical examples show, the new partial dual method enables us to find the best \( Z_I \)-type bounds within a much shorter period of time. At the same time this means that greater sized MDMPs can be bounded.
REFERENCES


