Journal of
Mathematical
Inequalities
Volume 3, Number 4 (2009), 485–498

INEQUALITIES FOR DISCRETE HIGHER ORDER CONVEX FUNCTIONS

ANDRÁS PRÉKOPA

Dedicated to Professor Josip Pečarić
on the occasion of his 60th birthday

1. Univariate Discrete Higher Order Convex Functions

Higher order convex functions have been first studied by Popoviciu (1934, 1944). There is a growing interest in this notion (see, e.g., Pečarić and Čuljak, 2002, Gilányi and Páles, 2008 and the references therein, among others), because these functions enjoy a number of good properties, important in applications.

The notion of higher order convexity is based on divided differences so much important in interpolation theory. Since divided differences can also be defined on discrete sets, the notion of a higher order convex function comes up in a natural way even in a purely mathematical framework.

However, the recent development in the theory of discrete higher order convex functions was necessitated by applications, primarily in reliability theory: bounding expectations and probabilities of Boolean functions of large numbers of events, where the calculation of the exact values are impossible, even by the use of contemporary hightech computers and calculations.

The paper where the notion of higher order discrete convexity is used, in combination with linear programming calculation, in a general framework, is the one by Prékopa (1990). In that paper a synthesis is given of earlier results of the same author and others and simultaneously the research area of the theory of discrete moment problems is initiated. The theory and the numerical calculations have already been extended to the multivariate case. This time, however, we restrict ourselves to the univariate case.

Let f be a function defined on the discrete set $z_0 < z_1 < \cdots < z_n$. The first order divided differences of f are defined by

$$[z_i, z_{i+1}]f = \frac{f(z_{i+1}) - f(z_i)}{z_{i+1} - z_i}, \qquad i = 0, 1, \ldots, n-1.$$

Mathematics subject classification (2000): 52A41.
Keywords and phrases: inequalities; higher order convex functions; expectation inequalities; power moment problem;.

Paper JMI-03-48

The kth order divided differences are defined recursively by

$$[z_i,\dots,z_{i+k}]f = \frac{]z_{i+1},\dots,z_{i+k}]f - [z_i,\dots,z_{i+k-1}]f}{z_{i+k}-z_i}, \qquad k \geqslant 2.$$

The function f is said to be kth order convex if all of its kth order divided differences are nonnegative. It is said to be kth order strictly convex, if all of its kth order divided differences are positive.

It is well known that (see, e.g., Jordan, 1947):

$$[z_i,\dots,z_{i+k}]f = \frac{\begin{vmatrix} 1 & 1 & \dots & 1 \\ z_i & z_{i+1} & \cdots & z_{i+k} \\ \vdots & \vdots & \ddots & \vdots \\ z_i^{k-1} & z_{i+1}^{k-1} & \dots & z_{i+k}^{k-1} \\ f(z_i) & f(z_{i+1}) & \dots & f(z_{i+k}) \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \dots & 1 \\ z_i & z_{i+1} & \cdots & z_{i+k} \\ \vdots & \vdots & \ddots & \vdots \\ z_i^{k-1} & z_{i+1}^{k-1} & \dots & z_{i+k}^{k-1} \\ z_i^k & z_{i+1}^k & \dots & z_{i+k}^k \end{vmatrix}}, \qquad 0 \leqslant i \leqslant n-k. \tag{1.1}$$

It is well known that (see, e.g., Popoviciu, 1944) the following theorem holds true.

THEOREM 1. *If the $(m+1)$st divided differences of the function f are positive on consecutive points, then all minors of order $m+2$ of the matrix*

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ z_0 & z_1 & \cdots & z_n \\ \vdots & \vdots & \ddots & \vdots \\ z_0^m & z_1^m & \dots & z_n^m \\ f(z_0) & f(z_1) & \dots & f(z_n) \end{pmatrix} \tag{1.2}$$

are positive. In other words, the $(m+1)$st divided differences corresponding to any $(m+1)$-element subset of $\{z_0,\dots,z_n\}$ are also positive.

In what follows we use extensively linear programming methodology. For an elementary introduction see Prékopa (1996).

Let ξ be a discrete random variable, the possible values of which are known to be the numbers $z_0 < z_1 < \cdots < z_n$. Introduce the notations

$$p_i = P(\xi = z_i), \qquad i = 0,1,\dots,n.$$

Suppose that the above probabilities are unknown but known are the power moments $\mu_k = E(\xi^k)$, $k = 1,\dots,m$, where $m < n$. Our aim is to minimize or maximize a linear functional, defined on $\{p_i\}$, subject to the constraints that arise from the moment equations. In other words, we consider the following linear programming problem:

$$\begin{array}{l}
\min\,(\max)\ \{f_0p_0+f_1p_1+\cdots+f_np_n\},\\
\text{subject to}\\
\quad p_0+\quad p_1\ +\cdots+\quad p_n=1,\\
z_0p_0+z_1p_1\ +\cdots+z_np_n=\mu_1,\\
z_0^2p_0+z_1^2p_1\ +\cdots+z_n^2p_n=\mu_2,\\
\qquad\vdots\\
z_0^mp_0+z_1^mp_1+\cdots+z_n^mp_n=\mu_m,\\
\quad p_0\geqslant 0,\ p_1\geqslant 0,\dots,p_n\geqslant 0.
\end{array} \tag{1.3}$$

In problem (1.3) the matrix A has full rank. Let B be an $(m+1)\times(m+1)$ part of A and designate by I or I_B the set of subscripts of those columns of A which form B. The collection of these vectors, as well as the matrix B, is called a basis. Sometimes we write $B(I)$ instead of B. Let f_B designate the vector of the basic components of f. The vector y satisfying

$$y^TB=f_B^T$$

is called the dual vector corresponding to B. The basis B is said to be dual feasible, relative to the minimization (maximization) problem, if we have

$$\begin{array}{ll}
y^Ta_p\leqslant f_p & \text{for } p\in\{0,\dots,n\}-I\\
(y^Ta_p\geqslant f_p & \text{for } p\in\{0,\dots,n\}-I).
\end{array} \tag{1.4}$$

If for every $p\in\{0,\dots,n\}-I$ we have $y^Ta_p\neq f_p$, then the basis is said to be dual nondegenerate.

The inequalities (1.4) are called the *condition of optimality* because if the basis B is primal feasible and (1.4) holds, then B is an optimal basis and the corresponding solution is an optimal solution to the problem. The differences $f_p-f_B^TB^{-1}a_p$ satisfy the equations

$$\begin{pmatrix}1 & f_B^T\\ 0 & B\end{pmatrix}\begin{pmatrix}f_p-f_B^TB^{-1}a_p\\ d_p\end{pmatrix}=\begin{pmatrix}f_p\\ a_p\end{pmatrix},$$

$p\in\{0,\dots,n\}-I$, hence we get the formulas

$$f_p-f_B^TB^{-1}a_p=\frac{\begin{vmatrix}f_p & f_B^T\\ a_p & B\end{vmatrix}}{\begin{vmatrix}1 & f_B^T\\ 0 & B\end{vmatrix}}=\frac{1}{|B|}\begin{vmatrix}f_p & f_B^T\\ a_p & B\end{vmatrix}, \tag{1.5}$$

$p\in\{0,\dots,n\}-I$.

Let $L_I(z)$ be the Lagrange polynomial of degree m, corresponding to the points z_i, $i\in I$, i.e.,

$$L_I(z)=\sum_{i\in I}f(z_i)L_{I,i}(z),$$

where

$$L_{I,i}(z) = \frac{\prod_{j\in I-\{i\}} (z-z_j)}{\prod_{j\in I-\{i\}} (z_i-z_j)}.$$

Define the vector

$$b(z) = \begin{pmatrix} 1 \\ z \\ \vdots \\ z^m \end{pmatrix}$$

for every real z. We assert that

$$f_B^T B^{-1}(I)b(z) = L_I(z). \tag{1.6}$$

In fact, $b(z_i) = a_i$ for $i \in I$, hence

$$f_B^T B^{-1}(I)b(z_i) = f(z_i), \qquad i \in I.$$

Thus, (1.6) holds for every real z.

¿From the above discussion a nice characterization follows, for the dual feasible bases, in terms of Lagrange polynomials: in the minimization (maximization) problem (1.3) a basis $B(I)$ is dual feasible if and only if the function $f(z)$ runs above (below) $L_I(z)$ for every z_i, $i \notin I$.

By a well-known formula in approximation theory, we have

$$f(z) - L_I(z) = \prod_{j\in I}(z-z_j)[z,z_i,\ i\in I]f, \tag{1.7}$$

valid for every z for which f is defined.

Any of the relations (1.5), (1.7) can be used to show that if we have knowledge about the sign of the divided differences $[z,z_i,\ i\in I]f$, then we can find out what subscript sets I determine dual feasible bases. This observation enables us to present simple proof for the following theorem.

THEOREM 2. *Suppose that all* $(m+1)$*st divided differences of the function* $f(z)$, $z \in \{z_0, z_1, \dots, z_n\}$ *are positive. Then, in problem* (1.3), *all bases are dual nondegenerate and the dual feasible bases have the following structures, presented in terms of the subscripts of the basic vectors:*

	$m+1$ *even*	$m+1$ *odd*
min problem	$\{j, j+1, \dots, k, k+1\}$	$\{0, j, j+1, \dots, k, k+1\}$
max problem	$\{0, j, j+1, \dots, k, k+1, n\}$	$\{j, j+1, \dots, k, k+1, n\}$,

where in all parentheses the numbers are arranged in increasing order.

1.1. Functional and Expectation Inequalities

If B is a dual feasible basis in the min problem (1.3), then

$$f(z) \geqslant L_I(z), \qquad I = I_B, \tag{1.8}$$

equality holds if $z \in I$ and the inequality is strict if $z \notin I$.

If B is a dual feasible basis in the max problem (1.3), then

$$f(z) \leqslant L_I(z), \qquad I = I_B, \tag{1.9}$$

equality holds if $z \in I$ and the inequality is strict if $z \notin I$.

It follows from (1.8) and (1.9) that under the given conditions we have

$$E(f(\xi)) \geqslant E(L_I(\xi)) \tag{1.10}$$

and

$$E(f(\xi)) \leqslant E(L_I(\xi)), \tag{1.11}$$

respectively. These inequalities are sharp, i.e., no better bounds can be given for $E(f(\xi))$ if we only know the first m moments of the distribution.

1.2. Special Case: Minimization Problem, $m+1=2$. Discrete Jensen's Inequality

The minimization problem (1.3) specializes as:

$$\begin{aligned}
&\min \{f_0p_0 + f_1p_1 + \cdots + f_np_n\}\\
&\text{subject to}\\
&\quad p_0 + \quad p_1 + \cdots + \quad p_n = 1\\
&z_0p_0 + z_1p_1 + \cdots + z_np_n = \mu\\
&\ p_0 \geqslant 0,\ p_1 \geqslant 0, \ldots, p_n \geqslant 0,
\end{aligned}$$

where $\mu = E(\xi)$ and the function f is second order convex. Any dual feasible basis has subscript set of the type $\{j, j+1\}$. If we solve for z_j, z_{j+1} the system of equations:

$$\begin{aligned}
p_j + \qquad p_{j+1} &= 1\\
z_jp_j + z_{j+1}p_{j+1} &= \mu,
\end{aligned}$$

then we obtain $p_j = (z_{j+1} - \mu)/(z_{j+1} - z_j)$, $p_{j+1} = 1 - p_j$,

$$E(f(\xi)) \geqslant f(z_j)\frac{z_{j+1} - \mu}{z_{j+1} - z_j} + f(z_{j+1})\frac{\mu - z_j}{z_{j+1} - z_j} \tag{1.12}$$

and the inequality is sharp if $p_j \geqslant 0$, $p_{j+1} \geqslant 0$, i.e., j is specified in such a way that $z_j \leqslant \mu \leqslant z_{j+1}$.

1.3. Special Case: Maximization Problem, $m+1=2$

The only dual feasible basis has subscript set $\{0,n\}$. ¿From the equation

$$\begin{aligned} p_0 + \quad p_n &= 1 \\ z_0 p_0 + z_n p_n &= \mu \end{aligned}$$

we obtain

$$p_0 = \frac{z_n - \mu}{z_n - z_0}, \qquad p_n = \frac{\mu - z_0}{z_n - z_0}$$

and the upper bound for $E(f(\xi))$ is expressed by the inequality:

$$E(f(\xi)) \leqslant f(z_0)\frac{z_n - \mu}{z_n - z_0} + f(z_n)\frac{\mu - z_0}{z_n - z_0}, \tag{1.13}$$

where f is second order convex. The inequality is sharp.

If $m+1=3$, then the third order divided differences of f have to be positive. The estimation of $E[f(\xi)]$ is based on the knowledge of μ_1 and μ_2. Since $m+1$ is odd, any dual feasible basis in the minimization (maximization) problem is of the form $\{0,i,i+1\}$ $(\{j,j+1,n\})$. We get the inequalities:

$$\begin{aligned} &\frac{z_i z_{i+1} - (z_i + z_{i+1})\mu_1 + \mu_2}{(z_i - z_0)(z_{i+1} - z_0)} f(z_0) - \frac{z_0 z_{i+1} - (z_0 + z_{i+1})\mu_1 + \mu_2}{(z_{i+1} - z_i)(z_i - z_0)} f(z_i) \\ &\quad + \frac{z_0 z_i - (z_0 + z_i)\mu_1 + \mu_2}{(z_{i+1} - z_i)(z_{i+1} - z_0)} f(z_{i+1}) \leqslant E(f(\xi)] \\ &\leqslant \frac{z_{j+1} z_n - (z_{j+1} + z_n)\mu_1 + \mu_2}{(z_{j+1} - z_j)(z_n - z_j)} f(z_j) - \frac{z_j z_n - (z_j + z_n)\mu_1 + \mu_2}{(z_{j+1} - z_j)(z_n - z_j)} f(z_{j+1}) \\ &\quad + \frac{z_j z_{j+1} - (z_j + z_{j+1})\mu_1 + \mu_2}{(z_{j+1} - z_j)(z_n - z_j)} f(z_n). \end{aligned} \tag{1.14}$$

The above inequalities are sharp if the bases are primal feasible too, i.e., i and j are determined by the inequalities

$$z_i \leqslant \frac{\mu_2 - z_0\mu_1}{\mu_1 - z_0} \leqslant z_{i+1}, \qquad z_j \leqslant \frac{z_n\mu_1 - \mu_2}{z_n - \mu_1} \leqslant z_{j+1}.$$

2. Bounds on the Probability of the Union of Events

Let $A_1,\ldots,A_n$ be arbitrary events in an arbitrary probability space. We want to compute or at least approximate the probability of $A_1 \cup \ldots \cup A_n$. A general method is provided by the inclusion–exclusion formula (see Takács 1967 for its history):

$$P(A_1 \cup \ldots \cup A_n) = S_1 - S_2 + \cdots + (-1)^{n-1} S_n, \tag{2.1}$$

where

$$S_k = \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} P(A_{i_1} \ldots A_{i_k}).$$

If n is large, then we may not be able to compute all S_k, $k=1,\ldots,n$ but only a few of them: $S_1,\ldots,S_m$, where $m<n$. We have the Bonferroni's bounds:

$$P(A_1\cup\ldots\cup A_n)\geqslant\sum_{k=1}^{m}(-1)^{k-1}S_k,\quad \text{if } m \text{ is even}$$
$$P(A_1\cup\ldots\cup A_n)\leqslant\sum_{k=1}^{m}(-1)^{k-1}S_k,\quad \text{if } m \text{ is odd.}$$

These bounds are usually weak. Fréchet (1940) has proved that the upper bound

$$P(A_1\cup\ldots\cup A_n)\leqslant S_1$$

is sharp. Dawson and Sankoff (1967) have proved that the sharp S_1, S_2 lower bound is:

$$P(A_1\cup\ldots\cup A_n)\geqslant\frac{2}{i+1}S_1-\frac{2}{i(i+1)}S_2,\qquad i=1+\left\lfloor\frac{2S_2}{S_1}\right\rfloor.$$

Kwerel (1975) has proved the sharp S_1, S_2 upper bound is:

$$P(A_1\cup\ldots\cup A_n)\leqslant S_1-\frac{2}{n}S_2.$$

THEOREM 3. *Let ν be the number of events (out of $A_1,\ldots,A_n$) that occur. We have the equation*

$$E\left[\binom{\nu}{k}\right]=S_k,\qquad k=1,\ldots,n,$$

i.e., S_k is the kth binomial moment of ν.

For a proof see, e.g., Prékopa (1995). The above equation holds true for $k=0$ if we define $S_0=1$.

The detailed form of Theorem 3 is:

$$\begin{aligned}
p_0+p_1+\ p_2+\ \ p_3+\cdots+\ \ p_n&=1\\
p_1+2p_2+\ 3p_3+\cdots+\ np_n&=S_1\\
p_2+\binom{3}{2}p_3+\cdots+\binom{n}{2}p_n&=S_2\\
\ddots\qquad\vdots\qquad&\\
\binom{n}{n}p_n&=S_n,
\end{aligned}\tag{2.2}$$

where $p_i=P(\nu=i)$, $i=0,\ldots,n$.

Given $p_0,\ldots,p_n$, we can determine $S_1,\ldots,S_n$ and vice versa.

Lower and upper bounds on the probability $P(\nu \geqslant 1) = P(A_1 \cup \ldots \cup A_n)$ can be obtained by the use of the optimum values of the linear programming problem:

$$\begin{aligned} &\min(\max) \sum_{i=1}^{n} p_i \\ &\text{subject to} \\ &\sum_{i=0}^{n} \binom{i}{k} p_i = S_k, \quad k = 0, \ldots, m \\ &\qquad\qquad p_i \geqslant 0, \qquad i = 0, \ldots, n. \end{aligned} \tag{2.3}$$

Suitable linear transformation on the equality constraints of problems (2.3) results in the power moment problem

$$\begin{aligned} &\min(\max) \sum_{i=1}^{n} p_i \\ &\text{subject to} \\ &\sum_{k=0}^{n} i^k p_i = \mu_k, \quad k = 0, \ldots, m \\ &\qquad\quad p_i \geqslant 0, \qquad i = 0, \ldots, n, \end{aligned} \tag{2.4}$$

where $\mu_k = E(\nu^k)$, $k = 0, \ldots, m$. The coefficient function in the objective of problem (2.4) is $m+1$-order convex (concave) if $m+1$ is odd (even).

A basis in problem (2.3) is primal (dual) feasible iff the corresponding basis in problem (2.4) is primal (dual) feasible. This fact enables us to handle the two problems simultaneously. The $S_1, \ldots, S_m$ binomial moments can be expressed by the use of the $\mu_1, \ldots, \mu_m$ power moments and vice versa.

Any formula or algorithm that provides us with the optimum value and optimal solution of problem (2.3), simultaneously provides us with those of problem (2.4).

Problem (2.3) was first used by Kwerel (1975), for the cases of $m = 2, 3$, to reproduce the S_1, S_2 sharp lower and upper bounds as well as produce new S_1, S_2, S_3 sharp lower and upper bounds.

A general theory of problems (2.3), (2.4) is due to Prékopa (1990).

To derive closed form or algorithmic bounds for the probability of the union, by the use of $S_1, \ldots, S_m$, it is more convenient to remove the first constraint and the first variable from problem (2.3). The new problem is:

$$\begin{aligned} &\min(\max) \sum_{i=1}^{n} p_i \\ &\text{subject to} \\ &\sum_{i=1}^{n} \binom{i}{k} p_i = S_k, \quad k = 1, \ldots, m \\ &\qquad\qquad p_i \geqslant 0, \qquad i = 1, \ldots, n. \end{aligned} \tag{2.5}$$

If $V_{\min}(V_{\max})$ designates the optimum value of the $\min(\max)$ problem (2.3), then the optimum value of the $\min(\max)$ problem (2.5) is $V_{\min}(\min(V_{\max}, 1))$. Below we present the sharp S_1, S_2, S_3 lower and upper bounds and the sharp S_1, S_2, S_3, S_4 upper bound.

The sharp S_1, S_2, S_3 lower bound (Kwerel 1975, Boros and Prékopa 1989) is:

$$\begin{gathered} P(A_1 \cup \ldots \cup A_n) \\ \geqslant \frac{i+2n-1}{(i+1)n} S_1 - \frac{2(2i+n-2)}{i(i+1)n} S_2 + \frac{6}{i(i+1)n} S_3 \\ i = 1 + \left\lfloor \frac{-6S_3 + 2(n-2)S_2}{-2S_2 + (n-1)S_1} \right\rfloor . \end{gathered}$$

The sharp S_1, S_2, S_3 upper bound (Kwerel 1975, Boros and Prékopa 1989) is:

$$\begin{gathered} P(A_1 \cup \ldots \cup A_n) \\ \leqslant \min\left(S_1 - \frac{2(2i-1)}{i(i+1)} S_2 + \frac{6}{i(i+1)} S_3, 1 \right) \\ i = 2 + \left\lfloor \frac{3S_3}{S_2} \right\rfloor . \end{gathered}$$

The sharp S_1, S_2, S_3, S_4 upper bound (Boros and Prékopa 1989) is:

$$\begin{gathered} P(A_1 \cup \ldots \cup A_n) \\ \leqslant \min\left(S_1 - \frac{2((i-1)(i-2)+(2i-1)n)}{i(i+1)n} S_2 + \frac{6(2i+n-4)}{i(i+1)n} S_3 - \frac{24}{i(i+1)n} S_4, 1 \right) \\ i = 1 + \left\lfloor \frac{-12S_4 + 3(n-4)S_3 + (n-2)S_2}{(n-2)S_2 - 3S_3} \right\rfloor . \end{gathered}$$

2.1. Algorithmic Solution of Problems (2.3)–(2.5)

First we present the dual feasible basis structure theorem (see Prékopa, 1988) that applies to the modified problem (2.5).

THEOREM 4. *All bases of problem* (2.5) *are dual nondegenerate and all dual feasible bases have the following structures, expressed in terms of the basic subscripts:*

	m even	*m odd*
min problem	$\{j, j+1, \ldots, k, k+1\}$	$\{j, j+1, \ldots, k, k+1, n\}$
max problem	$\{1, j, j+1, \ldots, k, k+1, n\}$	$\{1, j, j+1, \ldots, k, k+1\}$.

The algorithm can be presented in a unified manner, it applies simultaneously to problems (1.3), (2.3), (2.4), (2.5). It has the following steps.

Step 0. Find a dual feasible basis B by the use of the dual feasible basis structures.

Step 1. Check for $B^{-1}b \geqslant 0$. If yes, then go to Step 4. Otherwise go to Step 2.

Step 2. Identify j such that $(B^{-1}b)_j < 0$. Remove the jth vector from basis B.

Step 3. Include the unique vector, that restores dual feasible structure, into the basis. Go to Step 1.

Step 4. Stop, optimal basis and optimal solution have been obtained.

3. Improved Algorithm to Solve the Discrete Power Moment Problem

The algorithm is based on LU factorization of Vandermonde matrices. Even though Lagrange interpolation is fairly old, the above-mentioned LU factorization is more recent, see Turner (1966), Olver (2005) and the bibliography therein.

Let B be a basis in problem (1.3) and $I_B = \{i_0, \ldots, i_m\}$ the set of subscripts of the basic vectors. Below we present the general forms of the matrices L, U that provide us with the LU decomposition of the matrix B, and their inverses L^{-1}, U^{-1}. The first subscripts in L_{jk}, U_{jk}, $(L^{-1})_{jk}$, $(U^{-1})_{jk}$ indicate rows while the second subscripts indicate columns. We have the following formulas:

$$L_{jk} = \begin{cases} \displaystyle\sum_{\substack{l_1 \leqslant \cdots \leqslant l_{j-k} \\ \{l_1,\ldots,l_{j-k}\} \subset I_B}} z_{l_1} \ldots z_{l_{j-k}}, & \text{if } j \geqslant k \\ 0, & \text{if } j < k \end{cases} \tag{3.1}$$

$$(L^{-1})_{jk} = \begin{cases} (-1)^{j-k} \displaystyle\sum_{\substack{l_1 < \cdots < l_{j-k} \\ \{l_1,\ldots,l_{j-k}\} \subset I_B}} z_{l_1} \ldots z_{l_{j-k}}, & \text{if } j \geqslant k \\ 0, & \text{if } j < k \end{cases} \tag{3.2}$$

$$U_{jk} = \begin{cases} (z_{i_k} - z_{i_0})(z_{i_k} - z_{i_1}) \ldots (z_{i_k} - z_{i_{j-1}}), & \text{if } j \leqslant k \\ 0, & \text{if } j > k \end{cases} \tag{3.3}$$

$$(U^{-1})_{jk} = \begin{cases} (-1)^{k-j} \dfrac{1}{(z_{i_j} - z_{i_0})(z_{i_j} - z_{i_1}) \ldots (z_{i_j} - z_{i_{k-1}})}, & \text{if } j \geqslant k \\ 0, & \text{if } j < k, \end{cases} \tag{3.4}$$

$j, k = 0, \ldots, m$. In (3.1) the complete, in (3.2) the elementary symmetric functions are in the formulas.

We are interested only in the signs of the components of $B^{-1}b = U^{-1}L^{-1}b$. In view of the special forms of the entries of U^{-1} we can avoid division when we multiply $L^{-1}b$ by the rows of U^{-1}.

If we use the above LU decomposition for the bases of problem (1.3), then the algorithm of the previous section can be simplified and made more accurate. Let $b = (1, \mu, \ldots, \mu_m)^T$.

Step 0. Find an initial dual feasible basis B in agreement with the dual feasible basis structure theorem.

Step 1. Using the $B = LU$ decomposition, determine $L^{-1}b$ and begin to multiply $L^{-1}b$ by the rows of U^{-1}, starting with the first row. If we obtain nonnegative products for all rows go to Step 4. Otherwise go to Step 2.

Step 2. Let j be the subscript of the first row where we obtain negative product. Remove the jth vector from the basis, i.e., remove the jth base point from $\{z_i,\ i \in I_B\}$.

Step 3. Include that vector into the basis or point into the set of base points that restores dual feasibility. Go to Step 1.

Step 4. We have obtained $B^{-1}b = U^{-1}L^{-1}b \geqslant 0$ which means B is both primal and dual feasible, i.e. optimal.

4. Solution of the Continuous Power Moment Problem with Higher Order Convex Function in the Objective

Let $f(z)$, $a \leqslant z \leqslant b$ be a function such that all of its $m+1$ st order divided differences are positive and ξ a random variable for which $a \leqslant \xi \leqslant b$.

Suppose that the probability distribution of ξ is unknown but known are its first m moments $\mu_1, \ldots, \mu_m$. Let $F(z) = P(\xi \leqslant z)$ be the unknown c.d.f. of ξ, $-\infty < z < \infty$.

We want to solve the semi-infinite linear programming problem:

$$\begin{aligned} &\min(\max) \int_a^b f(z)\, \mathrm{d}F(z) \\ &\text{subject to} \\ &\int_a^b z^k\, \mathrm{d}F(z) = \mu_k, \qquad k = 0, \ldots, m. \end{aligned} \tag{4.1}$$

The dual of problem (4.1) is the following

$$\begin{aligned} &\max(\min) \sum_{k=0}^{m} y_k \mu_k \\ &\text{subject to} \\ &\sum_{k=0}^{m} y_k z^k \underset{(\geqslant)}{\leqslant} f(z), \qquad a \leqslant z \leqslant b. \end{aligned} \tag{4.2}$$

Problems of this and even more general type have been studied in the literature. The paper by Kemperman (1968) is very informative. Kemperman proved the duality theorem for problems (4.1), (4.2) and used the dual problem (4.2) to derive formulas for some special cases, where m is small.

Below we present an efficient iterative solution to problem (4.1). It is the combination of a cutting plane algorithm of Prékopa and Alexe (2003) and the algorithm of Prékopa presented in the previous section.

Step 0. Initiate $k \leftarrow 0$ and choose a finite grid $T_k = \{z_{k0}, \ldots, z_{kn_k}\}$ from the interval $[a, b]$.

Step 1. Write up problem (1.3) replacing T_k for $z_0, \ldots, z_n$ and using $f_i = f(z_{ki})$, $i = 0, \ldots, n_k$.

Step 2. Solve the problem by the dual algorithm presented in Section 4. Let B_k be the optimal basis.

Step 3. Construct the Lagrange polynomial

$$L_{B_k}(z) = f_{B_k}^T B_k^{-1} \begin{pmatrix} 1 \\ z \\ \vdots \\ z^m \end{pmatrix}.$$

We have the relation

$$f(z) - L_{B_k}(z) = [z_i,\ i \in I_{B_k},\ z; f] \prod_{i \in I_{B_k}} (z - z_i). \tag{4.3}$$

Since the function f has all positive divided differences in the interval $[a,b]$, it follows that

$$[z_{ki},\ i \in I_{B_k}, z; f] > 0 \text{ for every } z \notin \{z_{ki},\ i \in I_{B_k}\}. \tag{4.4}$$

Relation (4.3) and (4.4) imply that for any $z \notin \{z_{ki},\ i \in I_{B_k}\}$ we have

$$f(z) < L_{B_k}(z) \qquad (f(z) > L_{B_k}(z)),$$

between some consecutive basic points in the minimization (maximization) problem, otherwise the opposite inequalities hold, in agreement with the optimality criterion. Let

$$\begin{gathered} s = \operatorname*{argmin}_{z \in [a,b]} (f(z) - L_{B_k}(z)) \\ \left(s = \operatorname*{argmin}_{z \in [a,b]} (L_{B_k}(z) - f(z)) \right) \end{gathered} \tag{4.5}$$

in the minimization (maximization) problem. In view of the basis structure, the minimization in (4.5) may be restricted to the small subintervals of $[a,b]$ between some consecutive basic points (a great numerical advantage).

Given a tolerance limit $\varepsilon > 0$, the absolute value of the minimum in (4.5) is (a) smaller than or equal to ε, (b) greater than ε. In case (a) go to Step 4. In case of (b) supplement s to the grid, set $k \leftarrow k+1$ and go to step 1.

Step 4. Stop, the algorithm has terminated, the required precision in the solution of the problem has been obtained. The approximate optimal solution is the basic solution corresponding to B_k.

The correctness of the above algorithm follows from the correctness of the cutting plane method applied to the semi-infinite linear programming problem (see, e.g., Goberna, Lopez, 1998).

Prékopa and Szedmák (2002) report on the numerical results obtained in connection with the numerical solution of the discrete power moment problem. Moments of order up to 30 could be used in connection with a former algorithm. The use of *LU* decomposition can improve on it. Prékopa and Alexe (2003) report on the numerical solution of the continuous power moment problem, again, without the use of the *LU* decomposition. Problems with up to 30 moments could be solved with satisfactory precision. Again, we expect that the recent algorithm improves on the numerical solution and we will be able to solve even larger problems.

5. Relationship to Other Inequalities Involving Higher Order Convex Functions

Discrete and continuous moment problems are based on integrals

$$\mu_k = \int_\Omega \xi^k(\omega)\,\mathrm{d}P, \qquad k = 1,\dots,m$$

whereas most inequalities concerning higher order convex functions use derivatives.

An exception is the inequality

$$E(f(\xi)) \leqslant \frac{b-\mu}{b-a} f(a) + \frac{\mu - a}{b-a} f(b), \tag{5.1}$$

where ξ is a random variable such that $a \leqslant \xi \leqslant b$ and

$$\mu = E(\xi) = \int_\Omega \xi(\omega)\,\mathrm{d}P$$

is the expectation of ξ.

We can take, e.g., $\Omega = [c,d]$ and, designating the elements of Ω by x, rather than by ω, we can take an integrable function $g(x)$, $x \in [c,d]$ as a random variable. Assume that $a \leqslant g(x) \leqslant b$ for $c \leqslant x \leqslant d$. Then the inequality (5.1) takes the form

$$\int_a^b f(g(x))\,\mathrm{d}P(x) \leqslant \frac{b - \int_c^d g(x)\,\mathrm{d}P(x)}{b-a} f(a) + \frac{\int_c^d g(x)\,\mathrm{d}P(x) - a}{b-a} f(b),$$

where P is any nonnegative measure such that its value on $[c,d]$ is 1.

Inequality (5.1) is mentioned in Edmundson (1957) and Madansky (1959). They applied it to some problems in operations research.

REFERENCES

[1] E. BOROS AND A. PRÉKOPA, *Closed Form Two-Sided Bounds for Probabilities That Exactly r and at Least r out of n Events Occur*, Mathematics of Operations Research, **14** (1989), 317–342.

[2] D. DAWSON AND A. SANKOFF, *An Inequality for Probabilities*, Proceedings of the American Mathematical Society, **18** (1967), 504–507.

[3] H.P. EDMUNDSON, *Bounds on the Expectation of a Convex Function of a Random Variable*, The RAND Corporation, P–982, April 9, 1957.

[4] M. FRÉCHET, *Les Probabilités Associées à un Système d'Événement Compatibles et Dépendants, Actualités Scientifiques Industrielles*, Nos. 859, 942, Paris, 1940, 1943.

[5] A. GILÁNYI AND ZS. PÁLES, *On Convex Functions of Higher Order*, Math. Inequalities and Appl., **11** (2008), 271–282.

[6] A. GOBERNA AND M.A. LOPEZ, *Linear Semi-Infinite Optimization*, Wiley, New York, 1998.

[7] J.L. JENSEN, *Sur les fonctions convexes et les inégalités entre les valeurs moyennes*, Acta Math., **30** (1906), 175–193.

[8] C. JORDAN, *Calculus of Finite Differences*, Chelsea Publishing Company, New York, 1947.

[9] S. KARLIN AND W.J. STUDDEN, *Tchebycheff Systems: With Applications in Analysis and Statistics*, Interscience, New York, 1966.

[10] J.H.B. KEMPERMAN, *The General Moment Problem, a Geometric Approach*, Ann. Math. Statist., **39** (1968), 93–122.

[11] S.M. KWEREL, *Most Stringent Bounds on Aggregated Probabilities of Partially Specified Dependent Probability Systems*, Journal of the American Statistical Association, **70** (1975), 472–479.

[12] C.E. LEMKE, *The Dual Method for Solving the Linear Programming Problem*, Naval Research Logistic Quarterly, **1** (1954), 36–47.

[13] A. MADANSKY, *Bounds on the Expectation of a Convex Function of a Multivariate Random Variable*, Annals of Math. Stat., **30**, 743–746.

[14] P.J. OLVER, *On Multivariate Interpolation*, School of Math., Univ. of Minnesota, MN, 2005.

[15] J. PEČARIĆ AND V. ČULJAK, *Interpolation Polynomials and Convex Functions of Higher Order*, Math. Inequalities and Appl., **5** (2002), 369–386.

[16] T. POPOVICIU, *Sur quelques propriétés des fonctions d'une ou de deux variables réelles*, Mathematica (Cluj), **8** (1934), 1–85.

[17] T. POPOVICIU, *Les Fonctions Convexes. Actualités Scientifiques et Industrielles 992*, Hermann, Paris, France, 1944.

[18] A. PRÉKOPA, *Boole–Bonferroni Inequalities and Linear Programming*, Oper. Res., **36** (1988), 145–162.

[19] A. PRÉKOPA, *The Discrete Moment Problem and Linear Programming*, Discrete Applied Mathematics, **27** (1990), 235–254.

[20] A. PRÉKOPA, *Inequalities on Expectations Based on the Knowledge of Multivariate Moments*, M. Shaked, Y.L. Tong, eds., Stochastic Inequalities. Institute of Mathematical Statistics. Lecture Notes–Monograph Series, Vol. 22. Institute of Mathematical Statistics, Hayward, CA, 309–331, 1992.

[21] A. PRÉKOPA, *Stochastic Programming*, Kluwer Academic Publishers, Dordrecht, Boston, 1995.

[22] A. PRÉKOPA, *A Brief Introduction to Linear Programming*, Math. Scientist, **21** (1996), 85–111.

[23] A. PRÉKOPA, *On Multivariate Discrete Higher Order Convex Functions and Their Applications*, Proc. Sixth Internat. Conf. on Generalized Convexity and Monotonicity, August 29–September 2. Karlovasi, Samos, Greece, *Lecture Notes in Economics and Mathematical Systems*, Springer, Heidelberg, Germany, 21–47.

[24] A. PRÉKOPA AND G. ALEXE, *Dual Methods for the Numerical Solution of the Univariate Power Moment Problem*, RUTCOR Research Report **14**–2003.

[25] A. PRÉKOPA AND S. SZEDMÁK, *On the Numerical Solution of Discrete Power Moment Problem*, RUTCOR Research Report **99**–2002, Rutgers University, Piscataway, NJ.

[26] L. TAKÁCS, *On the Method of Inclusion and Exclusion*, Journal of the American Mathematical Association, **62** (1967), 102–113.

[27] L. RICHARD TURNER, *Inverse of the Vandermonde Matrix with Applications*, NASA Technical Note D-3547.

(Received July 30, 2008)

András Prékopa
RUTCOR, Rutgers Center for Operations Research
640 Bartholomew Road, Piscataway
NJ 08854–8003
e-mail: prekopa@rutcor.rutgers.edu