APPLICATION OF THE DISCRETE MOMENT PROBLEM FOR NUMERICAL INTEGRATION AND SOLUTION OF A SPECIAL TYPE OF MOMENT PROBLEMS

BY MARIYA NAUMOVA

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Operations Research

Written under the direction of

András Prékopa

and approved by

New Brunswick, New Jersey

October, 2015

ABSTRACT OF THE DISSERTATION

Application of the discrete moment problem for numerical integration and solution of a special type of moment problems

by Mariya Naumova

Dissertation Director: András Prékopa

We present a brief survey of some of the basic results related to the classical continuous moment problems (CMP) and the recently developed discrete moment problems (DMP), clarifying their relationship. We also introduce a new numerical integration method, based on DMP and termed Discrete Moment Method (DMM), that can be used for univariate piecewise higher order convex functions. This means that the interval where the function is defined can be subdivided into non-overlapping subintervals such that in each interval all divided differences of given orders, do not change the sign. The new method uses piecewise polynomial lower and upper bounds on the function, created in connection with suitable dual feasible bases in the univariate discrete moment problem and the integral of the function is approximated by tight lower and upper bounds on them. Numerical illustrations are presented for the cases of the normal, exponential, gamma and Weibull probability density functions. We show how a similar approach can be applied for solving the problems of a special structure, namely the discrete conditional moment problems and present the corresponding numerical results. Finally, we present novel applications to valuations of financial instruments.

Acknowledgements

First and foremost, I would like to thank my research advisor, Professor András Prékopa for his support, expertise and seemingly endless patience. I am deeply indebted to him for introducing me to this interesting research topic and for all that he has taught me throughout my time at Rutgers.

I would like to sincerely thank Dr. Andrzej Ruszczynski, Dr. Vladimir Gurvich, Dr. Melike Baykal-Gursoy, Dr. Myong K. Jeong, and Dr. Mine Subasi for agreeing to serve on my thesis committee, for their assistance and valuable comments.

I owe a special note of gratitude to Dr. Endre Boros for the interesting graduate courses I took from him and various forms of support during my graduate study.

In addition, I have been very privileged to get to know, to learn from and to collaborate with other brilliant scholars from Rutgers Center for Operations Research, such as Dr. Farid Alizadeh, Dr. Adi Ben-Israel, and Dr. Jonathan Eckstein.

Last but not least, I am and will always be thankful to my family and friends.

Table of Contents

Chapter 1

Introduction

The classical moment problem has received substantial attention since its inception in the second half of the XIX century. The major research work has been concentrated on two different types of problems dealing with

- bounding sums or integrals, based on the knowledge of a finite number of suitably defined moments;
- determination of the existence and uniqueness of a distribution such that its moments match with the elements of a finite or infinite sequence of real numbers.

The first problem was introduced by Bienaymé [7], Chebyshev [14] and Stieltjes [80]. While the second problem was formulated by Stieltjes [81] in connection with the continued fractions. Later on, the research done in regards with the second problem moved away from the continued fractions, and the research field was formed and explored, using various mathematical tools.

The term "Moment problem" itself appears for the first time in the work of Stieltjes [81].

Theory and numerical results in connection with the first type of the moment problems became important in many branches of mathematics and other sciences. They provide us with powerful methodologies to approximate sums and integrals, through bounding, given that a finite number of moments of a mass or probability distribution are numerically available.

Moment problems with discrete support for the mass distribution had been mentioned in the literature but they attracted great interest since the discovery of Prékopa [62], [64], [63] and Samuels, Studden [77] that the sharp Bonferroni bounds of Dawson and Sankoff [20], Sobel and Uppuluri [79] and others are optimum values of discrete binomial moment problems.

The structure of this dissertation is the following. In Chapter 2 we give a description of the discrete moment problem (DMP) and present a general solution method for it that is numerically stable. The proposed iterative algorithm solves a DMP formulated as a linear programming problem in the special case when the objective function has a property of being higher order convex. This means that the interval where the function is defined can be subdivided into non-overlapping subintervals such that in each interval all divided differences of given orders do not change the sign. The described methodology is based on Prékopa's approach to the DMP, where special linear programming formulations, theoretical results and new algorithms have been developed in the past twenty five years.

Some of the results presented here carry over to Chebyshev systems but in order to show the main ideas in a relatively simple way we restrict ourselves to the univariate power, and in part to the binomial moment problems. These problems, on the other hand, have remarkable connections to other fields of mathematics, such as interpolation and combinatorics. Some initial results in connection with discrete Chebyshev systems, using linear programming, are presented in [68].

In Chapter 3 we introduce a new numerical integration method, based on DMP, which is termed Discrete Moment Method (DMM). The method is designed for univariate functions that are piecewise higher order convex. The new method uses piecewise polynomial lower and upper bounds on the function, created in connection with suitable dual feasible bases in the univariate discrete moment problem and the integral of the function is approximated by tight lower and upper bounds on them. Numerical illustrations are presented for the cases of the normal, exponential, gamma and Weibull probability density functions.

In Chapter 4 we present a framework of solving discrete conditional moment problems using the methodology introduced in the previous chapters. The discrete conditional moment problems have a special "block" structure, and we present a modified Dantzig-Wolfe decomposition method for solving them.

In Chapter 5 we discuss how the discrete and continuous moment problems can be used in valuations of financial instruments and present some illustrative examples.

The conclusions are summarized in Chapter 6.

Chapter 2

Discrete Moment Problem

2.1 Problem formulation and related work

The famous "moment problem" was introduced in 1894 - 1895 by Stieltjes [81], [82], [53]. In his prominent work [81] he writes (see chapter 4):

Nous appellerons problème des moments le problème suivant :

Trouver une distribution de masse positive sur une droite $(0, \infty)$*,*

les moments d'ordre k *(*$k = 0, 1, ...$*) étant donnés.*

If μ_k denote these numbers, the problem consists in finding a positive measure σ on $[0, \infty)$ such that

$$\mu_k = \int_0^\infty x^k d\sigma(x), \quad k = 0, 1, ...$$

Stieltjes adopted the terminology "moment" from Mechanics and solved the problem using continued fractions. However, bounding problems related to moments had already been considered in works of Bienaymé [7] and Chebyshev [14], [15]. The bounding moment problem frequently appears in the literature as "Chebyshev type inequalities".

In 1884 Markov defended his doctoral thesis [57] where he demonstrated significant results for the bounding moment problems with the use of continued fractions, too. Later in 1919-1921, Hamburger [36], [37], [38] extended the Stieltjes moment problem to the real axis, and established the moment problem as a theory of its own.

In the same time, Hausdorff [39], [40] defined the Hausdorff moment problem on a finite interval in connection with convergence-preserving matrices; this new approach for the moment problem was the first one not related to continued fractions.

The Hamburger moment problem was extended to the complex functions by Nevanlinna [60]. He also provided solutions to the interpolation problem, now known as

Nevanlinna-Pick problem, that are related to the solutions of the associated power moment problem.

In 1923 Riesz [73] extended the moment problem in functional analysis, by observing the connection between the moment problem and the space of bounded linear functionals on $C([a,b])$. Around the same time, Carleman [11] showed the connection of the moment problem with the theories of quasi-analytic functions and of quadratic forms with infinitely many variables.

Several publications on the subject were done by Akhiezer and Krein [2], [3] who generalized the work by Markov. They assumed a finite number of moments in the problem. Riesz's theory was extended to the case of several dimensions by Haviland [41], [42] and Cramér [16].

Starting with the mid 1900's, the duality theory for the moment problem was developed independently by Isii [44], [45], and [47], in connection with the linear semi-infinite programming. Karlin [46], [47] explored from the geometric point of view a more general topic: the Chebyshev systems. However, the use of the duality theory for solving the bounding moment problem was proposed earlier in 1884 by Markov [57] and in 1911 by Riesz [72], [73].

Fundamental results in the duality theory for the moment problem were obtained by Haar[35], and Charnes, Cooper and Kortanek [12], [13]. In the mid 1900's, one of the most comprehensive studies dedicated to the use of the duality theory for solving the moment problem was written by Kemperman in 1968 [50].

At the end of 1980's, Prékopa [62], [64], [63] and Samuels and Studden [77] independently introduced and studied the univariate discrete moment problem, motivated by the fact that the sharp Bonferroni bounds, as well as other probability bounds, can be obtained as optimum values of discrete moment problems. Closed form formulas based on these results have been obtained by Boros and Prékopa in 1989 [9]. Few years later, Prékopa ([66], [67], [9]) introduced and studied the multivariate discrete moment problem. Although they address the same problem, the methodologies for solving the discrete moment problem used by Samuels and Studden [77], and Prékopa are completely different. Samuels and Studden use the classical approach for the general moment problem, and determine the solutions in closed form whenever possible; their method is applicable only to small size problems. Prékopa is the first who uses the linear programming methodology in moment theory, and it turns out that in the special case of the discrete moment problem, linear programming techniques provide us with more general and simpler algorithmic solutions than the classical ones. Moreover, the linear programming approach for the discrete moment problem allows for the efficient solution of solving efficiently large size moment problems, for which the classical methodology cannot give solutions, due to numerical difficulties.

2.2 Linear programming formulation of the univariate discrete moment problem

The material of this Section is based on [68] (Prékopa, 2001). Moment problems play an important role in many applied stochastic programming problems because we can obtain approximations of probabilities and expectations through the solutions of these problems. As it frequently happens, the probability distribution of a random variable is unknown, but a few moments of it are known and can be used to derive lower and upper bounds for the quantiles of the distribution and expectations of nonlinear (mostly convex) functions of the random variable.

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

$$p_i = P(\xi = z_i), \quad i = 0, 1, ..., n. \tag{2.2.1}$$

Suppose that probabilities $\{p_i\}$ are unknown but the power moments $\mu_k = E(\xi^k)$, $k = 1, ..., m$, where $m < n$ are known and let $\mu_0 = 1$.

Our aim is to find the upper and lower bounds for the expected value $\mathbb{E}[f(\xi)]$, which is possible by minimizing or maximizing 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 problems:

$$\begin{aligned} \min\ (\max)\ & \sum_{i=0}^{n} f(z_i) p_i \\ \text{subject to}\ & \sum_{i=0}^{n} z_i^k p_i = \mu_k,\ k = 0, \cdots, m, \\ & p_i \geq 0,\ i = 0, \cdots, n. \end{aligned} \tag{2.2.2}$$

The problem is called the discrete power moment problem. The set $\{p_i\}$ forms a probability distribution for the random variables ξ that has the support set $\{z_0, \cdots, z_n\}$.

Let the matrix of the equality constraints, its columns, and the right hand side vector be designated by $A, \mathbf{a}_0, \mathbf{a}_1, ..., \mathbf{a}_n$ and $\mathbf{b}$, respectively. Thus, we can rewrite the power moment problem in the following form:

$$\begin{aligned} &\min\ (\max)\ \sum_{i=0}^{n} f_i x_i \\ &\text{subject to}\ \sum_{i=0}^{n} \mathbf{a}_i x_i = \mathbf{b}, \\ &\qquad\qquad x_i \geq 0,\ \ i = 0, \cdots, n, \end{aligned} \tag{2.2.3}$$

where

$$\mathbf{a}_i = \begin{pmatrix} 1 \\ z_i \\ z_i^2 \\ \vdots \\ z_i^m \end{pmatrix}, \quad i = 0, 1, ..., n; \quad \mathbf{b} = \begin{pmatrix} 1 \\ \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{pmatrix}. \tag{2.2.4}$$

In the power moment problem the matrix $A = (\mathbf{a}_0, \mathbf{a}_1, ..., \mathbf{a}_n)$ is an $(m+1) \times (n+1)$ Vandermonde matrix. Hence every collection of $m + 1$ vectors of A forms a basis for the linear programming problem.

Below we present the background to describe the basis structure of the linear programming problem.

2.2.1 The properties of the divided differences and higher order convex functions

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

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

The kth order divided differences are defined recursively by

$$[z_i, ..., z_{i+k}]f = \frac{[z_{i+1}, ..., z_{i+k}]f - [z_i, ..., z_{i+k-1}]f}{z_{i+k} - z_i}, \quad k \geq 2.$$

Definition 2.2.1. *The function f is said to be kth order convex (strictly convex) if all of its kth order divided differences are nonnegative (positive).*

A sufficient condition for that is the following: f is defined in $[z_0, z_n]$ and has nonnegative (positive) kth order derivatives in $[z_0, z_n]$.

The first order convexity means that the function is nondecreasing, and the second order convexity means that the function is convex in the classical sense.

We can express the kth order divided differences by the following formula:

$$[z_i, ..., z_{i+k}]f = \frac{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ z_i & z_{i+1} & \cdots & z_{i+k} \\ \vdots & \vdots & \ddots & \vdots \\ z_i^{k-1} & z_{i+1}^{k-1} & \cdots & z_{i+k}^{k-1} \\ f(z_i) & f(z_{i+1}) & \cdots & f(z_{i+k}) \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ z_i & z_{i+1} & \cdots & z_{i+k} \\ \vdots & \vdots & \ddots & \vdots \\ z_i^{k-1} & z_{i+1}^{k-1} & \cdots & z_{i+k}^{k-1} \\ z_i^k & z_{i+1}^k & \cdots & z_{i+k}^k \end{vmatrix}}, \qquad 0 \leq i \leq n-k. \qquad (2.2.5)$$

The denominator in (2.2.5) is a Vandermonde determinant which is always positive, hence the sign of $[z_i, ..., z_{i+k}]f$ depends on the sign of the determinant standing in the numerator.

A Theorem (Prékopa, 1995) [67] states that if the kst order divided differences of the function f are positive on consecutive points, then all kst order divided differences of the function f are positive.

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

$$L_I(z) = \sum_{i \in I} f(z_i) L_{I,i}(z), \tag{2.2.6}$$

where

$$L_{I,i}(z) = \frac{\prod_{j \in I - \{i\}}(z - z_j)}{\prod_{j \in I - \{i\}}(z_i - z_j)}. \tag{2.2.7}$$

Define the vector

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

for every real z. We assert that

$$\mathbf{f}_B^T B^{-1}(I)\mathbf{b}(z) = L_I(z). \tag{2.2.8}$$

In fact, $\mathbf{b}(z_i) = \mathbf{a}_i$ for $i \in I$, hence

$$\mathbf{f}_B^T B^{-1}(I)\mathbf{b}(z_i) = f(z_i), \quad i \in I. \tag{2.2.9}$$

Thus, (2.2.8) holds for every real z. By a well-known formula in approximation theory, we have

$$f(z) - L_I(z) = \prod_{j \in I}(z - z_j)[\, z, z_i, i \in I\,]f, \tag{2.2.10}$$

valid for every z for which f is defined. From the above discussion a nice characterization follows, for the dual feasible bases, in terms of Lagrange polynomials: in the minimization (maximization) problem (2.2.2) 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, $i \notin I$.

2.2.2 The condition of the dual feasibility of the discrete power moment problem

The basis B is said to be primal feasible if

$$B^{-1}b \geq 0$$

and dual feasible in the minimization (maximization) problem if

$$f_i - f_B^T B^{-1} a_i \geq (\leq) 0,$$

$i = 0, 1, ..., n$.

Let B be a basis of A, and $I = \{i_0, \ldots, i_m\}$ the corresponding set of basic subscripts. B is called *dual feasible* if

$$f_B^t B^{-1} a_j \leq f_j, \quad \text{for every } j \in \{0, \ldots, n\} \setminus I, \tag{2.2.11}$$

in case of the minimization problem, and

$$f_B^t B^{-1} a_j \geq f_j, \quad \text{for every } j \in \{0, \ldots, n\} \setminus I, \tag{2.2.12}$$

in case of the maximization problem. B is called dual non-degenerate if $f_B^T B^{-1} a_j \neq f_j$, for every $j \in \{0, \ldots, n\} \setminus I$.

Theorem 2.2.1 (Prékopa 1990a)**.** *Suppose that all $(m+1)^{st}$ divided differences of the function f_z, $z \in \{z_0, \ldots, z_n\}$ are positive. Then in* (2.2.3) *all bases are dual nondegenerate and the dual feasible bases have the following structure:*

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

Proof. In Prékopa (1990b), two simple proofs are presented for the theorem. We recall one of them. Define the Lagrange polynomial

$$L_I(z) = \sum_{i=0}^{m} \frac{\prod\limits_{k \in I \setminus i} (z - z_k)}{\prod\limits_{k \in I \setminus i} (z_i - z_k)}.$$

It is well-known in interpolation theory that

$$f(z) - L_I(z) = \prod_{i \in I}(z - z_i)[z, z_i, i \in I; f] \tag{2.2.13}$$

holds for every $z \in \{z_0, \ldots, z_n\}$. Since $[z, z_i, i \in I; j] > 0$ for every $z \notin \{z_i, i \in I\}$, the assertion of the theorem follows from (2.2.13). $\square$

Thus, in case when the $(m+1)^{st}$ divided differences of f are positive, the dual feasible bases can be determined in closed form, with no computational effort. For this special case, Prékopa's dual algorithm for solving (2.2.3) can be applied as specified below (see Prékopa (2000) for a detailed presentation).

Prekopa's dual algorithm for the discrete moment problem

Step 1. Pick any dual feasible basis in agreement with the above result; let $I = \{i_0, \ldots, i_m\}$ be the set of basic indices.

Step 2. Determine the corresponding primal feasible solution $x_i = (B^{-1}b)_i$, for $i \in I$, and $x_i = 0$ for $i \in \{0, \ldots, n\} \setminus I$. If we take into account the formula:

$$\begin{vmatrix} a_0 & 1 & \ldots & 1 \\ a_1 & x_1 & \ldots & x_m \\ \ldots & \ldots & \ldots & \ldots \\ a_m & x_1^m & \ldots & x_m^m \end{vmatrix} = \begin{vmatrix} 1 & \ldots & 1 \\ x_1 & \ldots & x_m \\ \ldots & \ldots & \ldots \\ x_1^{m-1} & \ldots & x_m^{m-1} \end{vmatrix} \sum_{j=0}^{m} (-1)^j a_j S_{m-j}$$

where $S_j = \sum_{1 \le i_1 \le \cdots \le i_j \le m} z_{i_1} \ldots z_{i_j}$, $j = 0, \ldots, m$, we obtain that

$$x_{i_k} = \frac{\begin{vmatrix} \mu_0 & 1 & \ldots & 1 \\ \mu_1 & z_{i_0} & \ldots & z_{i_m} \\ \ldots & \ldots & \ldots & \ldots \\ \mu_m & z_{i_0}^m & \ldots & z_{i_m}^m \end{vmatrix}_{ik}}{\begin{vmatrix} 1 & 1 & \ldots & 1 \\ z_{i_k} & z_{i_0} & \ldots & z_{i_m} \\ \ldots & \ldots & \ldots & \ldots \\ z_{i_k} & z_{i_0}^m & \ldots & z_{i_m}^m \end{vmatrix}} = \frac{(-1)^{m-k} \sum_{j=0}^{m} (-1)^j \mu_j S_{m-j}}{\prod_{j=0}^{i_k - 1} (z_{i_k} - z_j) \prod_{j=i_k+1}^{i_m} (z_j - z_{i_k})},$$

for every $i_k \in I$.

- If $x_{i_k} \geq 0$, for every $i_k \in I$, then B is a primal-dual feasible basis, and therefore the current basic solution is optimal. Go to **Step 4**.
- If $x_{i_k} < 0$, for some i_k then the $i_k th$ vector of B is a candidate for outgoing. Go to **Step 3**.

Step 3. Include that vector into the basis that restores the dual feasible basis structure and go to **Step 2**.

Step 4. Stop. The optimal value $f_B^T B^{-1} b$ is a lower (upper) bound for $E[f(X)]$, depending on the type of the optimization problem (min or max, respectively).

The dual algorithm can also be applied in the more general case when the objective function f has nonnegative divided differences of order $m+1$, if some anti-cycling rule (e.g., lexicographic) is applied whenever dual degeneracy occurs.

An important property of the discrete power moment problems is that the optimal basis does not depend on the objective function as long as it has positive divided differences of order $m+1$. An immediate consequence of this fact is that the dual algorithm can be applied for the feasibility of the discrete power moment problem by taking an arbitrary objective function with positive divided differences of order $m+1$ (e.g., $f(z) = exp(z)$).

Due to the strong relationship between binomial and power moment problems, similar results can be obtained for the case of the binomial moment problem (see Prékopa (2000)).

Problem (2.2.3) can be used to find sharp lower and upper bounds for the probability of the union of n events: $A_1, \ldots, A_n$, where known are moments $\{\mu_k\}$ of the random variable ν, equal to the number of those events which occur. In this case the problem is

$$\begin{aligned} &\min(\max) && \sum_{i=1}^{n} x_i \\ &\text{subject to} && \\ & && \sum_{i=0}^{n} i^k x_i = \mu_k \quad k = 0, ..., m. \\ & && x_i \geq 0 \qquad\quad i = 0, ..., n, \end{aligned} \tag{2.2.14}$$

Let V_{min} (V_{max}) designate the optimum value of the min (max) problem. Problem (2.2.14) can equivalently be formulated by the use of the binomial moments $\{S_k\}_{k=0}^{m} = \{\binom{\nu}{k}\}_{k=0}^{m}$:

$$\begin{aligned} &\min(\max) && \sum_{i=1}^{n} x_i \\ &\text{subject to} && \\ & && \sum_{i=0}^{n} \binom{i}{k} x_i = \mu_k \quad k = 0, ..., m. \\ & && x_i \geq 0 \qquad\quad i = 0, ..., n, \end{aligned} \tag{2.2.15}$$

The optimum values are the same V_{min} (V_{max}), as in problem (2.2.14). Problem (2.2.14) and (2.2.15) can be simplified by removing variable x_0 and the only constraint that contains x_0. We obtain the problems:

$$\begin{aligned} &\min(\max) && \sum_{i=1}^{n} x_i \\ &\text{subject to} && \\ & && \sum_{i=1}^{n} i^k x_i = \mu_k \quad k = 0, ..., m. \\ & && x_i \geq 0 \qquad\quad i = 0, ..., n, \end{aligned} \tag{2.2.16}$$

and

$$\begin{aligned} &\min(\max) && \sum_{i=1}^{n} x_i \\ &\text{subject to} && \\ & && \sum_{i=1}^{n} \binom{i}{k} x_i = \mu_k \quad k = 0, ..., m. \\ & && x_i \geq 0 \qquad\quad i = 0, ..., n, \end{aligned} \tag{2.2.17}$$

Let W_{min} (W_{max}) designate the common optimum value of the minimization (maximization) problems (2.2.16) and (2.2.17). Prékopa (1990b) has shown that $V_{min} = W_{min}$

and $V_{max} = \min(W_{max}, 1)$. It is also shown in Prékopa (1990b) that Theorem 2.2.1 holds true for problems (2.2.16), (2.2.17).

Note that the discrete functions representing the objective function coefficients in the power moment problems (2.2.14) and (2.2.16) do not have all positive divided differences of order $m+1$ and m, respectively. The one in problem (2.2.14) has nonnegative divided differences if m is odd, and nonpositive divided differences if m is even. The one in problem (2.2.16), on the other hand, forms a Chebyshev system with the functions in the constraints, if m is even and its negative has the same property, if m is odd. Problems (2.2.16) and (2.2.17) have the same dual feasible basis structures as presented in Theorem 2.2.1, while (2.2.14) and (2.2.15) have somewhat different structures (see Prékopa (1990b)).

2.3 General Moment Problems

Let $\Omega = w$ with a σ-algebra S on it be a finite sequence of measurable functions $u_0(w), ..., u_m(w)$, $w \in \Omega$ and a finite sequence of real numbers $1, \mu_1, ..., \mu_m$, where $m \geq 1$. We consider the general moment problem that involves the following two problems:

- **The feasibility problem.** Find necessary and sufficient condition that μ_k is a generalized moment sequence with respect to the functions $u_k(w)$, i.e., there exists a measure P on S such that

$$\int_{\Omega} u_k(w)dP = \mu_k,\ k = 0, ..., m. \tag{2.3.1}$$

The system of equations (2.3.1) is called a feasible representation of the moment sequence μ_k.

- **The bounding problem.** Let $f(w)$, $w \in \Omega$ be a measurable function on S. Solve the optimization problem

$$\begin{aligned} &inf(sup) \qquad \textstyle\int_{\Omega} f(w)dP \\ &\text{subject to} \\ &\qquad\qquad \textstyle\int_{\Omega} u_k(w)dP = \mu_k, \\ &\qquad\qquad\qquad k = 0, ..., m. \end{aligned} \tag{2.3.2}$$

Let $P_{inf}(P_{sup})$ designate the optimum values of this problem.

The general moment problem is called determinate with respect to $u_k(w)$ if μ_k has a unique feasible representation (2.3.1), and indeterminate otherwise.

The general moment problem in some sense can be restricted to purely atomic probability measures, due to the following result of Richter ([71]), Rogosinksi ([76]) and Tchakaloff ([87]).

Theorem 2.3.1. *If the general moment problem is feasible, i.e, there exists a probability measure P such that* (2.3.1) *holds, then there exists a probability measure P, concentrated on at most $m + 1$ points, such that $\int_{\Omega} u_k(w)dP = \mu_k$, $k = 0, ..., m$.*

In most practical applications Ω is a subset of R_n, $n \geq 1$. If Ω is a compact real interval, the system of functions $u_k(z)$, $k = 0, ..., m$, is called a *Chebyshev system* of order m, if for all elements $z_0 < ... < z_m$ in Ω, the determinant

$$\begin{vmatrix} u_0(z_0) & \cdots & u_0(z_m) \\ \cdots & \cdots & \cdots \\ u_m(z_0) & \cdots & u_m(z_m) \end{vmatrix}. \tag{2.3.3}$$

is positive.

A Chebyshev system is called *weak* if the determinants are nonnegative, for any choice of $z_0, \cdots, z_m$.

Among the major references for Chebyshev systems we mention Karlin and Studden ([47]) and Krein and Nudelman ([58]).

We recognize that the matrix of the set of the constraints is a strongly ill-conditioned, and the computation of its inverse is a hard task even when the number of moments is relatively small. This computational difficulty and the restriction on the form of the function $f(z)$ claims a different approach to find the general solution of the moment problem. The concept based on the Chebyshev system renders the possibility to overcome these difficulties.

2.4 The numerical stability and the feasibility problem of the dual algorithm based on the Lagrangian approximation

The critical moment of the dual algorithm, which solves problem (2.2.3), is to recognize infeasibility of the problem that can occur in case of inaccurate moments. The inaccuracy can arise from measuring errors or rounding off. Such errors can cause infeasibility of problem (2.2.3) even in case when it is theoretically feasible. There are the following possibilities to prevent possible inaccuracy.

During the past two decades, numerous applications have arisen for high-precision floating-point arithmetic. High-precision floating-point arithmetic tools are standard features of Mathematica and Maple, and software packages such as

1. MPFR (available at http://www.mpfr.org/),
2. QD and ARPREC (available at http://crd-legacy.lbl.gov/ dhbailey/mpdist/).

More about high-precision floating-point arithmetic tools can be found in papers and presentations of Bailey from Lawrence Berkeley National Laboratory (see, for example, [5]) Some of these packages include high-level language interface modules that make conversion of standard-precision programs a relatively simple task. Applications of double-double (31 digits) or quad-double precision (62 digits) are particularly common, but there are also some interesting applications for as high as $50,000$ digits.

The price of the using the higher precision tools is the increasing computation time that can be many $(10-20)$ times more than the time required to perform a procedure with the embedded arithmetic tools.

Chapter 3

Numerical Integration

Numerical integration methods generally work in such a way that the integrand is evaluated at a finite number of points, called integration points or base points, and a weighted sum of these values approximates the integral. The base points and weights depend on the specific method used and the required accuracy.

An important part of the analysis of any numerical integration method is the study of the approximation error as a function of the number of integrand evaluations. A method which yields a small error for a small number of evaluations is usually considered efficient.

Numerical integration is widely used in statistical computation where it is applied in a straightforward way. Numerical integration methods serve as a tool in conversion of ordinary or partial differential equations into algebraic or variational equations. In boundary methods, numerical integration techniques are also used in transformations of partial differential equations into integral equations with subsequent discretization.

The use of numerical integration methods is not limited to mathematical applications, and the techniques of numerical integration have been a topic of an active interdisciplinary research. For example, numerical integration methods are used in evaluation of Bose-Einstein and Fermi-Dirac integrals that arise frequently in quantum statistics ([27], [52]).

The scientific literature is replete with the techniques of numerical integration, see for example Engels [22] and Davis and Rabinowitz [19] who list about a thousand papers on the topic. Many integration rules (see, e.g., [18], [43], [83]) use interpolation functions, typically polynomials, which are easy to integrate. The simplest rules of this type are the midpoint (or rectangle), the trapezoidal and the Simpson's rules, where for a small interval $[c, d]$ the approximations

$$\int_c^d f(x)dx \approx (d-c)f\left(\frac{c+d}{2}\right),$$

$$\int_c^d f(x)dx \approx (d-c)\left(\frac{f(c)+f(d)}{2}\right),$$

$$\int_c^d f(x)dx \approx \frac{d-c}{n}\left(\frac{f(c)+f(d)}{2} + \sum_{k=1}^{n-1} f\left(c + k\frac{d-c}{n}\right)\right),$$

respectively, are used. These methods rely on a "divide and conquer" strategy, whereby an integral on a relatively large set is broken down into integrals on smaller sets.

Interpolation with polynomials evaluated at equally-spaced points in $[c, d]$ yields the Newton-Cotes formulas, of which the rectangle and the trapezoidal rules are examples. Simpson's rule, which is based on a polynomial of order 2, is also a Newton-Cotes formula. If we allow the intervals between interpolation points to vary in length, we find other integration formulas, such as the Gaussian quadrature formulas. A Gaussian quadrature rule is typically more accurate than a Newton-Cotes rule which requires the same number of function evaluations, if the integrand is smooth. For a large number of variants of Gaussian quadrature the reader is referred to [31].

Romberg's method is based upon the approximation of the integral by the trapezoidal rule. Quadrature formulas of higher error order are produced by successive division of the step size by 2 and by an appropriate linear combination of the resulting approximations for the integral. First, one partitions $[c, d]$ into N_0 subintervals of length $h_0 = (c-d)/N_0$ and sets

$$N_i = 2^i N_0, h_i = h_0/2^i, i = 0, 1, ...,$$

then the integral is expressed as

$$\int_c^d f(x)dx = L_i^{(k)}(f(x)) + O(h_i^{2(k+1)}),$$

where $L_i^{(k)}(f(x))$ is a quadrature formula with error order $O(h_i^{2(k+1)})$. Romberg's method provides us with accurate results if the integrand has multiple continuous derivatives, though fairly good results may be obtained if only a few derivatives exist. We also mention numerical methods by [88]) that are useful when it is impossible or undesirable to use derivatives of the integrand.

The idea of the adaptive Simpson quadrature routine is described by Lyness [56], where the properties of the approximation error are also investigated. A recent discussion of error estimates and reliability of different adaptive routines based on Newton-Cote rules is given by Espelid [23]. The author tests available Matlab codes with respect to reliability and efficiency.

The literature on Gauss-Christoffel quadrature and its applications and computational applications is extensive. The famous method of approximate integration by the use of using a continued fraction expansion was discovered by Gauss in 1814, and throughout the 19th century it attracted the attention of many leading mathematicians of the time. For example, in 1826 Jacobi provided an alternative elegant derivation of the formula, showing that the nodes are the zeros of the Legendre polynomials and that they are real. Christoffel then significantly generalized the method and subsequently extended it to arbitrary measures of integration. The convergence of Gaussian quadrature methods was first studied by Stieltjes in 1884, while Markov endowed it with an error term. More on the history of the field of numerical integration can be found in Gautschi [26].

Thus, by the end of the 19th century, the integration method introduced by Gauss and Christoffel integration became widely known. It is, however, unlikely that it had actually been effectively used in practice, as the method requires the evaluation of functions at irrational arguments, and hence requires some tedious interpolation. The 20th century, with the development of powerful digital tools for computing, brought in a renewed interest in Gauss-Christoffel quadrature, and the formulae began to be frequently applied, which in turn, led to important new theoretical developments.

The importance of the eigenvalues and eigenvectors of the Jacobi matrices for computing Gauss' quadrature rules was demonstrated by Golub and Welsch [32], and the generalization to Radau and Lobatto quadrature was described by Golub and Kautsky [30].

A significant survey of different numerical integration methods was done by Dahlquist [18], [17], Gautschi [25], [26], and Mysovskih [59]. Gautschi [28] also described the use orthogonal polynomials for approximation purposes and Gauss-Christoffel quadrature computation.

Laurie [54] provides a survey on approximation methods and discusses some practical error estimation techniques in numerical integration. He underlines the difficulty of the formal theoretical error estimates for quadrature rules. Many Gaussian quadrature formulas with various weight functions are tabulated in Stroud and Secrest [85].

Recently, some algorithms and computer codes for generating integration rules became available in the public domain. Kautsky and Elhay [48] have developed algorithms and a collection of Fortran subroutines called IQPACK [49] for computing weights of interpolatory quadratures. In [61] R. Piessens et al. describe QUADPACK, a collection of Fortran 77 and 90 subroutines for numerical integration. The authors provide descriptions of the algorithms, program listings, test programs and examples. They also include useful advice on numerical integration and many references to the numerical integration literature used in developing QUADPACK. The CQUAD integration algorithm is described in the paper by Gonnet [33].

A package QPQ developed by Gautschi consisting of MATLAB programs for generating orthogonal polynomials as well as dealing with applications is available in public domain at www.cs.purdue.edu/archives/2002/wxg/codes. The package includes routines for generating Gauss-type arbitrary weight functions.

A set of Maple programs for Gauss quadrature rules are given by von Matt [89]. Gaussian product rules for integration over the n-dimensional cube, sphere, surface of a sphere, and tetrahedron are derived in Stroud and Secrest [85], and some simple formulas of various accuracy are tabulated in [1]. The derivation of such formulas are treated by Engels [22]. Nonproduct rules for multidimensional integration are found in Stroud [84].

Genz has developed a database of algorithms for numerical computation of multiple integrals available at http://www.math.wsu.edu/faculty/genz/homepage. These integrals arise in a wide variety application areas that include electromagnetics, chemistry, physics and statistics. He also presented a survey and new results in multivariate approximation theory [29].

In the following sections we propose a new univariate numerical integration method.

We create lower and upper bounding polynomials for the function on a finite grid but ensure that the integrals of the bounding polynomials provide us with tight lower and upper bounds for the integral of our function in an entire interval. We use Lagrange polynomials for bounding that are natural outcomes of the use of the discrete power moment problem. We illustrate our new method for the functions: $e^{x^2/2}$, $x^m e^{-x^2/2}$, $\left(\frac{x}{\lambda}\right)^{m-1} e^{-\left(\frac{x}{\lambda}\right)^m}$, and $\lambda e^{-\lambda x}$.

3.1 Conditions on the Base Points to Obtain Bounds on the Integral

Let f be a convex function of order $m+1$ in the interval $[c,d]$ and $Z = \{z_0, z_1, ..., z_n\} \subset [a,b]$. Suppose that the set $\{z_{i_0}, z_{i_1}, ..., z_{i_m}\} \subset Z$ defines a dual feasible basis in minimization problem (2.2.2) and let $l(z)$ designate the corresponding Lagrange polynomial (for simplicity we suppress the subscript B). We have the relation

$$f(z) - l(z) = [z_{i_0}, z_{i_1}, ..., z_{i_m}, z; f] \prod_{k=0}^{m} (z - z_{i_k}) \geq 0 \tag{3.1.1}$$

for any $z \in Z$. If it is dual feasible in the maximization problem (2.2.2) and the corresponding Lagrange polynomial is $u(z)$, then we have the relation:

$$f(z) - u(z) = [z_{i_0}, z_{i_1}, ..., z_{i_m}, z; f] \prod_{k=0}^{m} (z - z_{i_k}) \leq 0 \tag{3.1.2}$$

for any $z \in Z$. In both cases equality holds for $z \in \{z_{i_0}, z_{i_1}, ..., z_{i_m}\}$. Inequalities (3.1.1) and (3.1.2) hold true also for $z \in [a,b]$ with the exception of the interiors of consecutive pairs, described in Prékopa's Dual Theorem, among the base points $\{z_{i_0}, z_{i_1}, ..., z_{i_m}\}$, where the inequalities are reversed. For this reason from (3.1.1) and (3.1.2) we cannot immediately derive that

$$\int_c^d l(z)dz \leq \int_c^d f(z)dz \leq \int_c^d u(z)dz \tag{3.1.3}$$

However, the intervals between the consecutive pairs are small and in practice there is a relatively small number of consecutive pairs, hence the integrals of the differences $f(z) - l(z)$, $u(z) - f(z)$ over the union of consecutive pairs are small and allow for the validity of the relations in (3.1.3). Figures 3.1 and 3.2 illustrate the situation.

In Figure (3.1) the graphs show that if the base points $z \in \{z_0, z_j, z_{j+1}, z_k, z_{k+1}\}$ are chosen in such a way that z_j, z_{j+1} as well as z_k, z_{k+1} are close to each other, then $l(z) \geq f(z)$ on the small intervals $[z_j, z_{j+1}]$, $[z_k, z_{k+1}]$, otherwise we have $l(z) \leq f(z)$. The deficiency in the integral $\int_a^b l(z)dz$ caused by $l(z) \geq f(z)$ in $(z_j, z_{j+1}) \cup (z_k, z_{k+1})$ can easily be offset by choosing in a suitable way. The same idea applies to the maximization problem (see Fig.3.2).

Stating it in a different way: under mild conditions on the function and the base

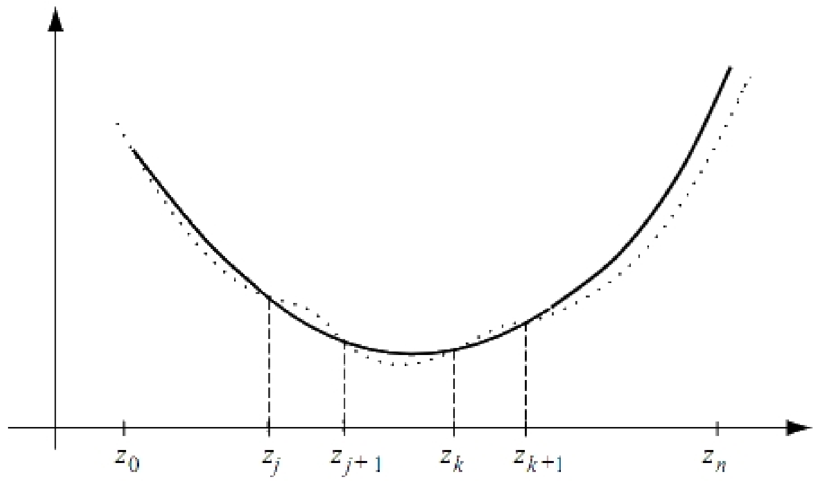

Figure 3.1: Function $f(z)$; Lagrange polynomial (dotted line); minimization problem; $m+1$ odd; basic subscript set: $\{z_j, z_{j+1}, z_k, z_{k+1}, z_n\}$

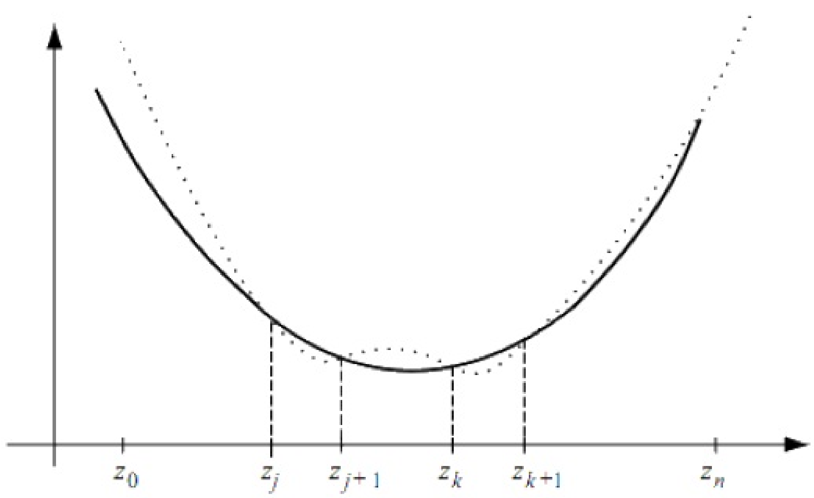

Figure 3.2: Function $f(z)$; Lagrange polynomial (dotted line); maximization problem; $m+1$ odd; basic subscript set: $\{z_0, z_j, z_{j+1}, z_k, z_{k+1}\}$

points the nonpositivity of the integrals of $f(z) - l(z)$ over the intervals between consecutive pairs is offset by the nonnegativity of the integrals over the much larger set, where $f(z) - l(z) \geq 0$, and a similar assertion holds for $u(z) - f(z)$.

This ensures that we in fact have the relations (3.1.3). But we can say more about it. If that happens then the fact that in some intervals the integral of $f(z) - l(z)$ is negative makes the lower bound tighter and the negativity of $u(z) - f(z)$ in some intervals makes the upper bound tighter.

3.2 Some strategies to find suitable base points to prove inequalities (3.1.3)

The subintervals of the interval $[c, d]$, where $f(z) - l(z) < 0$, $u(z) - f(z) < 0$, are small and the effect of the negativity of the integrals of these differences can easily be offset in several ways, under some conditions. We choose subintervals, where $f(z) - l(z) < 0$, $u(z) - f(z) < 0$, at one of the ends of $[c, d]$.

Lemma 3.2.1. *Let $[u, v]$ be an interval with positive length, then for any $y \in [u, v]$ and $z \le \underline{z} = \frac{u+v}{2} - \frac{\sqrt{2}}{2}(v - u)$ we have inequality*

$$(y - u)(v - y) \le (z - u)(z - v). \tag{3.2.1}$$

Proof. The largest value on the left-hand side is $\frac{(v-u)^2}{4}$. The right-hand side is increasing if $z \ge v$ and decreasing if $z \le u$. For $z = \bar{z}$ and $z = \underline{z}$ it is equal to $\frac{(v-u)^2}{4}$.

Theorem 3.2.1. *Let $f(z)$, $a \le z \le b$ be a real valued function that has nonnegative divided differences of orders $m + 1$.*

I. Suppose that $m + 1$ is even and the function has nonnegative divided differences of order $m + 2$. Take $\frac{m+1}{2}$ equidistant subintervals from $[a, b]$: $[u_1, v_1]$, $[u_2, v_2]$, $\dots$, $[u_{(m+1)/2}, v_{(m+1)/2}]$, $a < u_1 < v_1 < \cdots < u_{(m+1)/2} < v_{(m+1)/2} < b$, such that

$$\sum_{i=1}^{(m+1)/2} (v_i - u_i) \le b - \bar{z}, \tag{3.2.2}$$

where $\bar{z} = \frac{u_{(m+1)/2} + v_{(m+1)/2}}{2} + \frac{\sqrt{2}}{2}(v_{(m+1)/2} - u_{(m+1)/2})$.

Let $L(z)$ be the Lagrange polynomial corresponding to the base points u_j, v_j, $j = 1, \dots, \frac{m+1}{2}$.

Under these conditions we have the inequality

$$-\sum_{j=1}^{(m+1)/2} \int_{u_j}^{v_j} (f(z) - L(z))dz \le \int_{b-\bar{z}}^{b} (f(z) - L(z))dz. \tag{3.2.3}$$

If $m + 1$ is odd then we take $m/2$ subintervals and create the Lagrange polynomial by the use of all endpoints, supplemented by the left-hand endpoint of the interval $[a, b]$. Relation (3.2.3) remains true, if we replace m for $m + 1$ in (3.2.1), (3.2.2), (3.2.3).

II. Suppose that $m+1$ is even and the function has non-positive divided differences of order $m+2$.

Take $\frac{m+1}{2}$ subintervals from $[a,b]$, as in Case I, but assume

$$\sum_{j=1}^{(m+1)/2} (v_j - u_j) \leq \underline{z} - a, \tag{3.2.4}$$

where $\underline{z} = \frac{u_1+v_1}{2} - \frac{\sqrt{2}}{2}(v_1 - u_1)$.

Under these conditions we have the inequality:

$$-\sum_{j=1}^{(m+1)/2} \int_{u_j}^{v_j} (f(z) - L(z))dz \leq \int_a^{\bar{z}} (f(z) - L(z))dz. \tag{3.2.5}$$

Proof. We prove the assertion in Case I, for $m+1$ even. The proofs of the other assertions are the same. Assume that the columns of A, corresponding to u_j, v_j, $j = 1, ..., \frac{m+1}{2}$, form a dual feasible basis in the minimization problem (2.2.3). Then the Lagrange polynomial with the same base points has the property that

$$f(z) - L(z) \geq 0, \text{ if } z \in \bigcup_{j=1}^{\frac{m+1}{2}} [u_j, v_j]. \tag{3.2.6}$$

Equality holds if and only if z is one of the points u_j, v_j, $j = 1, ..., \frac{m+1}{2}$. We offset the negative integral of $L(z) - f(z)$ over the union of the intervals $[u_j, v_j]$, $j = 1, ..., \frac{m+1}{2}$, by the use of the integral of $f(z) - L(z)$ over the interval $[b - \bar{z}, b]$. This will be accomplished if

$$\begin{aligned} &-\left[u_i, v_i, i = 1, ..., \frac{m+1}{2}, y; f\right] \prod_{j=1}^{\frac{m+1}{2}} (y - u_j)(y - v_j) \\ &\leq \left[u_i, v_i, i = 1, ..., \frac{m+1}{2}, z; f\right] \prod_{j=1}^{\frac{m+1}{2}} (z - u_j)(z - v_j), \end{aligned} \tag{3.2.7}$$

for any $y \in \bigcup_{j=1}^{\frac{m+1}{2}} [u_j, v_j]$ and $z \geq \bar{z}$.

To show (3.2.7), first, we remark that the non-negativity of the divided differences of order $m+2$ of the function f implies the inequality:

$$-\left[u_i, v_i, i = 1, ..., \frac{m+1}{2}, y; f\right] \leq \left[u_i, v_i, i = 1, ..., \frac{m+1}{2}, z; f\right]. \tag{3.2.8}$$

Hence we only have to prove that for any y and z, satisfying the above condition, we have

$$-\prod_{j=1}^{\frac{m+1}{2}}(y-u_j)(y-v_j) \leq \prod_{j=1}^{\frac{m+1}{2}}(z-u_j)(z-v_j). \tag{3.2.9}$$

3.3 Strategies to improve on the accuracy of the integration

If X is a uniformly distributed random variable in $[a, b]$, then

$$E(f(X)) = \frac{1}{b-a} \int_a^b f(z)dz. \tag{3.3.1}$$

On the other hand, the uniform distribution in a finite interval is determined by its moment sequence $\mu_1, \mu_2, \ldots$, which is a special case of the Hausdorff one-dimensional moment problem (see, [40] and [78]). Thus, if $\mu_1, \mu_2, \ldots$ are the moments of the uniform distribution in $[a, b]$, then the optimum values of the linear programming problems

$$\begin{aligned} &\min\ (\max) \int_a^b f(z)dP \\ &\text{subject to } z^k dP = \mu_k \\ &\qquad\qquad k = 1, \ldots, m+1, \end{aligned} \tag{3.3.2}$$

provide us with lower and upper bounds for the integral (3.3.1). The two bounds converge to the integral (3.3.1) if $m \to \infty$.

Problem (**??**) can be solved approximately by the use of a sufficiently fine discretization of the interval. We can go from the discrete optimum to the continuous optimum by the use of Prékopa's dual algorithm. In [69], for example, the authors use the same approach providing a method for solving the continuous power moment problem when some higher order divided differences of the objective function are nonnegative. Their method combines Prékopa's dual approach for solving the discrete moment problem with a cutting-plane type procedure for solving linear semi-infinite programming problems.

We may not want to accurately compute the integral (**??**), instead, we may want to increase the number of constraints. In our numerical integration technique this means that if we work with a fine grid (a fine discretization of the interval $[a, b]$), and pick a subset of the grid point that determine a dual feasible basis of the discrete moment problem, then we may increase the number of elements of this subset in agreement with the dual feasible basis structure theorem.

For example, if $m+1$ is even, and we have $\frac{m+1}{2}$ consecutive pairs of the grid points, in a minimization problem, then we may include into the set of selected grid points a further one, to have $m+2$ points altogether. This will provide us with a dual feasible basis for the discrete moment problem with $m+2$ constraints if and only if the new grid point is the left hand endpoint of the set of grid points.

As regards error bounds for our integration technique, since the integral is approximated by integrals of Lagrange polynomials, we can use the error bounds available for Lagrange interpolation.

Our integration technique works for piece-wise higher order convex function, so we may restrict ourselves to one such piece of the function. Let $[a,b]$ be the interval where the piece is defined. We also assume that the function f is has continuous derivatives of order $m+1$ in that interval. It is not a serious restriction of generality, from the practical point of view. In fact, we can further subdivide the interval $[a,b]$, to ensure that the above differentiability condition holds true, at least in typical practical situation.

Let $z_0, z_1, ..., z_n$ be the set of base points for the lower bounding Lagrange polynomial $L(z)$ and $h = \max_i (z_{i+1} - z_i)$. Then the error bound for the lower bounding polynomial $L(z)$ is given by

$$\max_z \left(f(z) - L(z)\right) \leq \frac{f^{m+1}(\xi) h^{m+1}}{4(m+1)}, \tag{3.3.3}$$

where ξ is defined in such a way that the divided difference f of order $m+1$ at the base points of the Lagrange polynomial is equal to $f^{m+1}(\xi)$.

Similar inequality holds for the upper bounding polynomial. The base points chosen for $U(z)$ are different, hence instead of ξ and h we have other values η and k.

The error bound is

$$\max_z \left(U(z) - f(z)\right) \leq \frac{f^{m+1}(\eta) k^{m+1}}{4(m+1)}. \tag{3.3.4}$$

Based on (3.3.3) and (3.3.4) we can set up error bounds for the approximate integrals of $f(z)$ on $[a,b]$ we have inequalities:

$$\left| \int_a^b \left(f(z) - L(z)\right) dz \right| \leq \int_a^b \left| \left(f(z) - L(z)\right) \right| dz \leq \frac{f^{m+1}(\xi) h^{m+1}}{4(m+1)} \left(b - a - \alpha\right), \tag{3.3.5}$$

$$\left| \int_a^b \left(U(z) - f(z)\right) dz \right| \leq \int_a^b \left| \left(U(z) - f(z)\right) \right| dz \leq \frac{f^{m+1}(\eta) k^{m+1}}{4(m+1)} \left(b - a - \beta\right). \tag{3.3.6}$$

3.4 The Discrete Moment Method (DMM) of Univariate Numerical Integration

In this section we briefly describe the new numerical integration method, which is one of the main contributions of this work.

If a function $f(z)$, $a \leq z \leq b$, is convex of order $m+1$, then for any discrete set of points Z of at least $m+2$ points the discretized function $f(z)$ has all nonnegative divided differences of order $m+1$. As we have seen, we can construct two m-degree polynomials $l(z)$ and $u(z)$ such that

$$l(z) \leq f(z) \leq u(z), \; z \in Z. \tag{3.4.1}$$

The bounding polynomials can be obtained by the use of dual feasible bases, corresponding to problem (2.2.3). Then we approximate

$$\int_a^b f(z)dz \tag{3.4.2}$$

by the integrals

$$\int_a^b l(z)dz, \int_a^b u(z)dz. \tag{3.4.3}$$

Our intention is not only to approximate the integral (3.4.2) by the integrals (3.4.3) but to ensure that the integrals (3.4.3) serve as lower and upper bounds, respectively for the integral (3.4.2), i.e.,

$$\int_a^b l(z)dz \leq \int_a^b f(z)dz \leq \int_a^b u(z)dz. \tag{3.4.4}$$

As we have mentioned in the previous section, inequality (3.4.4) is not a direct consequence of (3.4.1) but its validity can be ensured by suitable choices of the points in the base sets, in other words, the dual feasible bases in problem (2.2.3) that define the polynomials $l(z)$, $u(z)$.

The function $f(z)$ may not be higher order convex (or concave) in the entire interval $[a, b]$. However, it may be true that the interval $[a, b]$ can be subdivided into a finite number of non-overlapping intervals such that on each of them the function is higher (not necessarily always of the same) order convex (concave). If this is the case, then we apply the numerical integration procedure for each subdividing interval and create bounds and approximations of the integral of $f(z)$ on $[a, b]$ by the use of the integrals on the subdividing intervals.

Algorithm

Initialization:

Use Lsum as the notation that the lower bound summation of integral of Lagrange polynomials in subintervals;

Use Usum as the notation that the upper bound summation of integral of Lagrange polynomials in subintervals.

Procedure:

Step 1. Determine the subdividing intervals where the rth divided difference of the function is positive or negative.

Step 2. For each subdividing interval $[c, d]$

a) If the rth divided difference of the function is positive, repeat - subdivide it into n subintervals of equal length $(n \geq r)$;

- label those endpoints by $z_0, z_1, ..., z_n$, evaluate the function at these labeled points;

- find any dual feasible basis according to Theorem 1 to get its corresponding upper and lower bounding Lagrange polynomials;

- integrate the upper and lower bounding Lagrange polynomials in the subinterval;

- Lsum = Lsum + Integral of the lower bounding Lagrange polynomial in $[c, d]$;

- Usum = Usum + Integral of the upper bounding Lagrange polynomial in $[c, d]$.

b) If the rth divided difference of the function is negative, multiply the function by -1, continue as in part a.

Note that the error in this new numerical integration method can easily be controlled because we provide simultaneous lower and upper bounds for the integral. If the bounds are not close enough then we may increase the number of base points to increase accuracy. The inclusion of new base points increases the lower bound and decreases the upper bound.

3.5 Illustration for the Case of the Normal Probability Density Function

We pay special attention to the univariate normal probability density function because of the connection to orthogonal polynomials.

The Hermite polynomials, a classical sequence of orthogonal polynomials, arise e.g. in probability theory, combinatorics and physics, and can be defined as

$$H_r(x) = (-1)^r e^{\frac{x^2}{2}} \frac{d^r}{dx^r} \left(e^{\frac{-x^2}{2}} \right), \tag{3.5.1}$$

or as

$$\tilde{H}_r(x) = (-1)^r e^{x^2} \frac{d^r}{dx^r} \left(e^{-x^2} \right), \tag{3.5.2}$$

hese two definitions are not exactly equivalent; either is a rescaling of the other, more precisely,

$$\tilde{H}_r(x) = 2^{\frac{r}{2}} H_r\left(\sqrt{2}x\right). \tag{3.5.3}$$

We use the first definition which is often preferred in probabilistic applications. In fact, $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ is the probability density function of the standard normal distribution. The first ten Hermite polynomials are:

$$H_0(x) = 1;$$

$$H_1(x) = x;$$

$$H_2(x) = x^2 - 1;$$

$$H_3(x) = x^3 - 3x;$$

$$H_4(x) = x^4 - 6x^2 + 3;$$

$$H_5(x) = x^5 - 10x^3 + 15x;$$

$$H_6(x) = x^6 - 15x^4 + 45x^2 - 15;$$

$$H_7(x) = x^7 - 21x^5 + 105x^3 - 105x;$$

$$H_8(x) = x^8 - 28x^6 + 210x^4 - 420x^2 + 105;$$

$$H_9(x) = x^9 - 36x^7 + 378x^5 - 1260x^3 + 945x.$$

The roots of the Hermite polynomials for $r = 2$ to $r = 10$ have been tabulated to eight decimals and are presented in the table below. Because of symmetry it is enough

Table 3.1: The roots of $H_r(x)$ for $r = 2$ to $r = 10$

r	Roots
2	1.00000000
3	0.00000000
	1.73205081
4	0.74196378
	2.33441422
5	0.00000000
	1.35562618
	2.85697001
6	0.61670659
	1.88917588
	3.32425743
7	0.00000000
	1.15440539
	2.36675941
	3.75043971
8	0.53907981
	1.63651904
	2.80248586
	4.14454719
9	0.00000000
	1.02325566
	2.07684798
	3.20542900
	4.51274586
10	0.48493571
	1.46600182
	2.48432584
	3.58182348
	4.85946283

to present the nonnegative values.

Once the roots of $H_r(x)$ are found, it is possible to determine the intervals where $\frac{d^r}{dx^r}\left(e^{\frac{-x^2}{2}}\right)$ is positive or negative. Therefore, for each interval, if $\frac{d^r}{dx^r}\left(e^{\frac{-x^2}{2}}\right)$ is positive, the lower (upper) bound of $e^{\frac{-x^2}{2}}$ is the value at x of the Lagrange polynomial, associated with the minimization (maximization) problem (2.2.3), where $f(x) = e^{\frac{-x^2}{2}}$. If $\frac{d^r}{dx^r}\left(e^{\frac{-x^2}{2}}\right)$ is negative, the lower (upper) bound of $e^{\frac{-x^2}{2}}$ is the value at x of the Lagrange polynomial, associated with the minimization (maximization) problem.

Hence, we propose the following algorithm to approximate the normal integral in interval $[a, b]$.

Algorithm

Step 1. Calculate the rth roots of the Hermite polynomial in the interval $[a, b]$.

Step 2. Determine the subdividing intervals where the rth derivative of $f(x) = e^{\frac{-x^2}{2}}$ is positive or negative.

Step 3. For each subdividing interval $[c, d]$,

a) If the rth derivative of $f(x) = e^{\frac{-x^2}{2}}$ is positive, repeat

- subdivide it into n subintervals of equal length $(n \geq d)$;

- label those endpoints by $z_0, z_1, ..., z_n$, evaluate the function at these labeled points and construct the following problem:

$$\begin{aligned} &\min\ (\max)\ \sum_{i=0}^{n} f(z_i)p_i \\ &\text{subject to } \sum_{i=0}^{n} z_i^k p_i = \mu_k,\ k = 0, \cdots, m, \\ &\qquad\qquad p_i \geq 0,\ i = 0, \cdots, n, \end{aligned} \tag{3.5.4}$$

that is problem (2.2.2).

- Find any dual feasible basis according to Theorem 1 to get its corresponding upper and lower bounding Lagrange polynomials.

- Integrate the upper and lower bounding Lagrange polynomials in the subinterval.

- Lsum = Lsum + Integral of the lower bounding Lagrange polynomial in $[c, d]$.

- Usum = Usum + Integral of the upper bounding Lagrange polynomial in $[c, d]$.

b) If the rth derivative of $f(x) = e^{\frac{-x^2}{2}}$ is negative, multiply the function by -1, continue as in part a.

Step 4. Multiply the Lsum and Usum by $\frac{1}{\sqrt{2\pi}}$ to get the lower and upper bound for the integral of univariate normal in interval $[a, b]$.

3.6 Further Numerical Results

We evaluated the probability integrals of the following functions: $e^{x^2/2}$, $x^m e^{-x^2/2}$, $\left(\frac{x}{\lambda}\right)^{m-1} e^{-\left(\frac{x}{\lambda}\right)^m}$, and $\lambda e^{-\lambda x}$ with different parameters in (a, b).

For a fixed r we consider the zeros of the Hermite polynomials that are in the interval (a, b). Each interval between two zeroes (or between one zero and one endpoint of (a, b)), we divide into k subintervals, and each subinterval - into N smaller intervals of equal length. For each small interval we choose M points to generate two Lagrange polynomials. Integration of these polynomials and summation over all the subintervals yields the final bounds.

The results for different functions and parameters are presented in the tables below.

Table 3.2: $f(x) = e^{-\frac{x^2}{2}}, a = 0, b = 2$

M	k	N	$Lsum$	$Usum$	Average	Exact value (8 digits of accuracy)
3	6	40	1.19513897	1.19754021	1.19633959	1.19628801
3	7	40	1.19552001	1.19711352	1.19631676	1.19628801
3	8	40	1.19583924	1.19682013	1.19632969	1.19628801
3	10	40	1.19604298	1.19655469	1.19629884	1.19628801
3	15	40	1.19621297	1.19638119	1.19629708	1.19628801
4	5	40	1.19627345	1.19631213	1.19629279	1.19628801
5	5	40	1.19628863	1.19628999	1.19628969	1.19628801

Table 3.3: $f(x) = e^{-\frac{x^2}{2}}, a = 0, b = 3$

M	k	N	$Lsum$	$Usum$	Average	Exact value (8 digits of accuracy)
3	6	40	1.24872355	1.25018107	1.24945231	1.24993045
3	7	40	1.24917761	1.25028238	1.24972910	1.24993045
3	8	40	1.24942248	1.25016039	1.24979144	1.24993045
3	10	40	1.24967086	1.25005136	1.24986111	1.24993045
3	15	40	1.24985429	1.24997429	1.24991429	1.24993045
4	5	40	1.24991898	1.24995206	1.24993552	1.24993045
5	5	40	1.24993042	1.24993235	1.24993139	1.24993045
5	10	40	1.24993046	1.24993079	1.24993063	1.24993045

Table 3.4: $f(x) = x^m e^{-\frac{x^2}{2}}, m = 3, a = 0, b = 2$

M	k	N	$Lsum$	$Usum$	Average	Exact value (5 digits of accuracy)
3	6	20	1.18644	1.18930	1.18787	1.18799
3	7	20	1.18692	1.18889	1.18791	1.18799
3	8	20	1.18723	1.18858	1.18791	1.18799
3	10	20	1.18755	1.18832	1.18792	1.18799
3	15	20	1.18779	1.18813	1.18796	1.18799
4	5	20	1.18795	1.18806	1.18801	1.18799
5	5	20	1.18798	1.18801	1.18800	1.18799
5	10	20	1.18799	1.18798	1.18799	1.18799

Table 3.5: $f(x) = x^m e^{-\frac{x^2}{2}}, m = 3, a = 0, b = 4$

M	k	N	$Lsum$	$Usum$	Average	Exact value (5 digits of accuracy)
3	6	20	1.99258	1.99539	1.99399	1.99396
3	7	20	1.99302	1.99489	1.99396	1.99396
3	8	20	1.99334	1.99453	1.99394	1.99396
3	10	20	1.99361	1.99425	1.99393	1.99396
3	15	20	1.99380	1.99407	1.99394	1.99396
4	5	20	1.99387	1.99402	1.99395	1.99396
5	5	20	1.99393	1.99398	1.99396	1.99396
5	10	20	1.99395	1.99397	1.99396	1.99396

Table 3.6: $f(x) = (\frac{x}{\lambda})^{m-1} e^{-(\frac{x}{\lambda})^m}, m = 3, a = 0, b = 4$

M	k	N	$Lsum$	$Usum$	Average	Exact value (5 digits of accuracy)
3	6	20	0.84150	0.84410	0.84280	0.84283
3	7	20	0.84194	0.84368	0.84281	0.84283
3	8	20	0.84225	0.84336	0.84281	0.84283
3	10	20	0.84253	0.84311	0.84282	0.84283
3	15	20	0.84271	0.84293	0.84282	0.84283
4	5	20	0.84276	0.84287	0.84282	0.84283
5	5	20	0.84280	0.84284	0.84282	0.84283
5	10	20	0.84282	0.84283	0.84283	0.84283

Table 3.7: $f(x) = (\frac{x}{\lambda})^{m-1} e^{-(\frac{x}{\lambda})^m}, m = 3, a = 0, b = 4$

M	k	N	$Lsum$	$Usum$	Average	Exact value (5 digits of accuracy)
3	6	20	0.86439	0.86479	0.86459	0.86466
3	7	20	0.86444	0.86471	0.86458	0.86466
3	8	20	0.86465	0.86466	0.86466	0.86466

Chapter 4

Discrete Conditional Moment Problem

4.1 Motivation

DMP and BMP can be solved efficiently using Prékopa's dual method up to 30 moments for the three types of functions:

1. Higher order convex;

2. Constants up to a point, then higher order convex;

3. Constant at only one point and 0 otherwise.

Numerical experiments show that for functions of type II (when the function is discontinuous and the jump is big) and type III, although the bounds are sharp, the quality of the lower and upper bounds are not enough for practical application. Moreover, there is no stable algorithm for other types of objective function (aside from three special types of functions above) due to the lack of the dual feasible basis structure. Therefore, we are facing the following problems:

1. Can we find the optimal lower and upper bounds for a larger variety of objective functions?

2. Can we develop a fast and stable algorithm to handle problems with even higher moments?

3. How to incorporate more information into the moment problems?

To address the above questions, we propose a new approach for the problem of bounding the expectation of a function of a discrete random variable for the case of piecewise higher order convex functions. The bounds are based on the knowledge of some of the power moments as well as conditional moments. The discrete conditional moment bounding problems (DCMP) are formulated as LPs with special structures and can be solved using Dantzig-Wolfe decomposition by the use of the Discrete Moment Problem (DMP). This brings more information into the moment problems in two ways: (1) we include conditional moments; (2) we include information about the shape of the distribution.

We contribute to the theory of discrete moment problems in many ways. First, the new formulation helps to incorporate more moments, more information about the distribution into the discrete moment problems. Second, we can solve a much large scale problem due to the use of decomposition principle. Third, the method is more stable since we deal with one subproblem at a time and if we use a small number of conditional moments, the subproblems can be solved in closed forms. Last but not least, we can find optimal bounds for a much larger class of functions. For example, if the objective function is higher order convex of order 3, then we can use up to 2 conditional moments constraints for the subproblem.

In the following section we present the formulation and the decomposition principle for bounding expectation of a random variable by the use of the dual feasible basis structure theorems.

4.2 Problem formulation

The discrete conditional moment problem can be formulated as a linear programming problem in the following way:

$$\begin{array}{llllll}
\min(\max) & g_0^T p_0 & +g_1^T p_1 & +\dots & +g_n^T p_n & \\
\text{subject to} & & & & & \\
 & M_0 p_0 & +\dots & & M_l p_l & = b_0 \\
 & A_0 p_0 & & & & = b_1 \\
 & \vdots & & \ddots & & = \vdots \\
 & & & & A_l p_l & = b_l \\
 & p_0 \geq 0, & \dots & & , p_l \geq 0, &
\end{array} \tag{4.2.1}$$

where M_i is an $(1+m_0) \times n_i$ matrix, A_i is an $m_i \times n_i$ matrix, $(M_0, ..., M_n)$ is Vandermonde, A_i's are Vandermonde matrix of size $(m_i+1) \times n_i$ with the same numbers in its columns as in m_i but different powers, $p_i^T = (p_{i_1}, ..., p_{i_{n_i}})$, $i = 1, \ldots, l$, the components of g_i form a higher order convex function, $i = 1, \ldots, l, m_i < m_0$. $b_1, ..., b_l$ are vectors of conditional moments.

Here $M_0, ..., M_l$ and $A_0, ..., A_l$ can be chosen in a flexible manner. For practical application, it may be better to have $A_0, ..., A_l$ as matrices corresponding to conditional moments of lower order (e.g., from 1 to k) since they can be easily computed (not very expensive computation). Higher moments can be included in $M_0, ..., M_l$. In this case, M_i has typical columns: $[1\ z_j^k\ \ldots\ z_i^{k+m_0}]^T$ and $f_i^T M_i$ means an $k+m_0-$degree polynomial. It is higher order convex or concave depending on the sign of the last component of f_i.

There is no problem, however, if we choose a special matrix $M_i = [1 \ldots 1]$ (hence, $b_0 = 1$), i.e., we do not use any full moments in the formulation.

$$
\begin{array}{llllll}
\min(\max) & g_0^T p_0 & +g_1^T p_1 & +\ldots & +g_n^T p_n & \\
\text{subject to} & & & & & \\
 & p_0 & +\ldots & & p_l & = 1 \\
 & A_0 p_0 & & & & = b_1 \\
 & \vdots & & \ddots & & = \vdots \\
 & & & & A_l p_l & = b_l \\
 & p_0 \geq 0, & \ldots & & , p_l \geq 0 &
\end{array}
\tag{4.2.2}
$$

Let B be a feasible basis in the master problem and consider two types of subproblems. The first type is a discrete moment problem given some moments of the distribution.

$$
\begin{array}{ll}
\max & \sum_{i=0}^{n} (f_i^T M_i - g_i^T) p_i \\
\text{subject to} & \\
 & A_i p_i = b_i \\
 & p_i \geq 0,
\end{array}
$$

where f_i are the dual variables associated with equations 1 to $1+m_0$ in the master problem. Under the formulation (4.2.2), problem (4.2.3) takes a simpler forms

$$\begin{aligned}
\max \quad & \textstyle\sum_{i=0}^{n}(f_i - g_i^T)p_i \\
\text{subject to} \quad & \\
& A_i p_i = b_i \\
& p_i \geq 0,
\end{aligned}$$

where f_i is just a constant. Hence, the subproblems can be handled by DMP, not the master problem, however. In this case, the original problem can be min or max.

Dual feasible basis structures for (4.2.3) has extensively been studied in Prékopa (1988, 1989, 1990a, 1990b, 2008, 2009). They can be solved fast and efficiently by the use of Prékopa's dual method.

The feasible solution of a subproblem can be found by applying Prékopa's dual method for the subproblem on a higher order convex function and we take the optimal solution as one of the column for the decomposition.

4.3 The discrete conditional moment problems with given shapes of the distribution

If we know that the distribution is unimodal or multimodal, and we either know the exact location of the mode or the mode is given in an interval, then we can also incorporate these information. In this case, the objective function has to be high-order convex (or piece-wise higher order convex at the cutpoints).

The second type of the subproblem can include one more constraint on the increasing (decreasing) property of p_i. In this case, the right hand side b_i are conditional moments around the modes.

$$\begin{aligned}
\max \quad & \textstyle\sum_{i=0}^{n}(f_i^T M_i - g_i^T)p_i \\
\text{subject to} \quad & \\
& A_i p_i = b_i \\
& p_i < p_{i+1} < \cdots < p_l
\end{aligned}$$

When A_i is a matrix of binomial moments, then it can be transformed into a matrix of power moment. Thus, problem (4.2.3) can be solved using Prékopa's dual method for the discrete power moment problem.

Dual feasible basis structures for (4.3.1) has recently been studied in Prékopa (2008, 2009). They can be solved efficiently by the use of the special structure of the dual feasible basis.

4.4 The Dantzig-Wolfe decomposition method applied to the discrete conditional moment problem

Consider the following linear programming problem with the "block" constraints:

$$\begin{aligned} \min \quad & \textstyle\sum_{k=1}^t c_k^T x_k \\ \text{subject to} \quad & \\ & \textstyle\sum_{k=1}^t M_k x_k = b_0 \\ & A_k x_k = b_k, \ \ k = 1, ..., t \\ & x_k \geq 0, \ \ k = 1, ..., t \end{aligned} \tag{4.4.1}$$

We represent the solutions satisfying constraints

$$A_k x_k = b_k x_k \geq 0, \tag{4.4.2}$$

i.e., this convex polyhedron, in the following form

$$x_k = \sum_{i=1}^{r_k} \lambda_{ki} p_{ki} + \sum_{i=1}^{s_k} \mu_{ki} q_{ki}, k = 1, ..., t, \tag{4.4.3}$$

where

$$\lambda_{ki} \geq 0, i = 1, ..., r_k, \sum_{i=1}^{r_k} \lambda_{ki} = 1,$$
$$\mu_{ki} \geq 0, i = 1, ..., s_k, k = 1, ..., t.$$

If we replace vectors x_k in problem (4.4.1), and introduce the notations

$$\begin{aligned} M_k p_{ki} &= g_{ki}, \\ M_k q_{ki} &= h_{ki}, \\ c_k^T p_{ki} &= u_{ki}, \\ c_k^T q_{ki} &= v_{ki}, \end{aligned} \tag{4.4.4}$$

the problem takes the form

$$
\begin{array}{ll}
\min & \sum_{k=1}^{t} (\sum_{i=1}^{r_k} \lambda_{ki} u_i + \sum_{i=1}^{s_k} \mu_{ki} v_{ki}) \\
\text{subject to} &
\end{array}
$$

$$
\begin{gathered}
\sum_{k=1}^{t} (\sum_{i=1}^{r_k} g_{ki} \lambda_{ki} + \sum_{i=1}^{s_k} h_{ki} \mu_{ki}) = b_0 \\
\sum_{i=1}^{r_k} \lambda_{ki} = 1 \\
\lambda_{ki} \geq 0, i = 1, ..., r_k, \\
\mu_{ki} \geq 0, i = 1, ..., s_k, k = 1, ..., t.
\end{gathered}
\tag{4.4.5}
$$

Problem (4.4.5) that we obtain by modifying the original problem is often called the *master problem.*

Assume that A_k is of size $m_k \times n_k$ and M_k is of size $m_0 \times n_k$. Then in problem (4.4.1), there are $m_0 + m_1 + ... + m_n$ equality constraints and $n_1 + ... + n_t$ variables, while in problem (4.4.5) there is a large number of variables, but only $m_0 + t$ equality constraints.

A set of constraints of the original problem (4.4.1) is identified as "connecting" or "coupling" constraints, wherein many of the variables contained in the constraints have non-zero coefficients. The remaining constraints are to be grouped into independent submatrices such that if a variable has a non-zero coefficient within one submatrix, it will not have a non-zero coefficient in another submatrix.

The structure of the matrices of these two problems is shown below.

Table 4.1: The structure of (4.4.1)

M_1	M_2		M_t
A_1			
	A_2		
		$\ddots$	
			A_t

Assume that the bases concerning problem (4.4.5) are quadratic and that we know one feasible basis that is denoted B. Typically to the notation in the revised simplex method, we denote the vector consisting of the coefficients, corresponding to the

Table 4.2: The structure of (4.4.5)

m_0	g_{11}	...	g_{1r_1}	h_{11}	...	h_{1s_1}	...	g_{t1}	...	g_{1r_1}	h_{11}	...	h_{1s_1}
1	1	...	1	0	...	0							
$\vdots$							$\ddots$						
t								1	...	1	0	...	0

columns of B in the objective function, c_B. Let us partition vector $c_B^T B^{-1}$ as follows:

$$c_B^T B^{-1} = \left(f_1^T, f_2^T\right), \tag{4.4.6}$$

where f_1 has m components and f_2 has t components.

Then, we analyze the quantities

$$\begin{aligned} z_{ki}^{(g)} - u_{ki} &= f_1^T g_{ki} + f_{2k} - u_{ki}, \\ & i = 1, ..., r_k, \\ z_{ki}^{(h)} - u_{ki} &= f_1^T h_{ki} - v_{ki}, \\ & i = 1, ..., s_k,\ k = 1, ..., t, \end{aligned} \tag{4.4.7}$$

where f_{2k} is the kth component of f_2. If all the quantities in (4.4.7) are nonpositive, then B is the optimal basis. These quantities, however, are not known explicitly, except for those which belong to the basis. Therefore, we can overcome the difficulty by using the auxiliary subproblems:

$$\begin{aligned} \max \quad & \left(f_1^T M_k - c_k^T\right) x_k \\ \text{subject to} \quad & \\ & A_k x_k = b_k, \\ & x_k \geq 0. \end{aligned} \tag{4.4.8}$$

In fact, if we are using the (lexicographic) simplex method for the solution of problem (4.4.8), the method provides us with an extreme point of the convex polyhedron determined by the constraints, if there exists a finite optimum. Assuming this to be the case and that the optimal solution is p_{ki}, for the optimum value we obtain

$$\left(f_1^T M_k - c_k^T\right) p_{ki} = f_1^T g_{ki} - u_{ki} = z_{ki}^g - u_{ki} - f_{2k}. \tag{4.4.9}$$

Since f_{2k} is unknown, from here we can obtain $z_{ki}^g - u_{ki}$.

If, for every value k, the convex polyhedron defined by (4.4.2) is bounded, then there are no h_{ki} vectors and no numbers in the second row of (4.4.7). On the other hand, in this case all subproblems (4.4.8) have finite optima.

We solve all subproblems (4.4.8) and check if the execution of the algorithm has terminated so that all differences in the first row of (4.4.7) are nonpositive. If this is not the case, we have to continue the process. In the latter case, the vector $\begin{pmatrix} g_{ki} \\ e_k \end{pmatrix}$ may enter the basis, where $g_{ki} = M_k p_{ki}$ and p_{ki} is the optimal solution of subproblem (4.4.8), provided that f_{2k} plus the optimum value of this subproblem (equal to $z_{ki}^{(g)} - u_{ki}$) is positive. The determination of the vector leaving the basis and the updating of the revised simplex method tableau goes in the usual way.

Dropping the condition regarding the boundedness of the convex polyhedra (4.4.2), an iteration remains the same whenever all subproblems have finite optima. If, however, one of the subproblems (4.4.8) does not have a finite optimum, then we can find an extremal ray of the corresponding convex polyhedron (4.4.2) in the following way. Take the dual tableau, without the first row, corresponding to the last basis, when solving problem (4.4.8) and take that column in it which produced the information that there is no finite optimum. (We solve problem (4.4.8) by the simplex method and use the dual tableau only at the end.) The negative of that column is an extremal ray of the convex polyhedron (4.4.8). This must be among the vectors standing in the second sum in the representation (4.4.3). Let q_{ki} designate this vector. Since the scalar product of g_{ki} and the coefficient vector of the objective function is positive (by construction of q_{ki}),

$$\left(f_1^T M_k - e_k^T\right) q_{ki} = f_1^T h_{ki} - v_{ki} > 0, \tag{4.4.10}$$

which means that $q_k i$ may enter the basis. The lexicographic revised method can be used to guarantee finiteness. The initial basis can be found by the first phase of the two phase method. When we solve Master problem (4.4.5), the vectors (4.4.3) constitute the optimal solution of problem (4.4.1), where the $\lambda_{ki}, \mu_{ki}, p_{ki}, q_{ki}$ are taken from the final tableau, corresponding to the master problem (4.4.5).

4.5 Application of the Dantzig-Wolfe decomposition to the conditional moment problem

We know that if $(M_1 M_2 ... M_t)$ in (4.4.1) is a Vandermonde matrix, the individually $M_1, M_2, ..., M_t$ are Vandermonde matrices, and so are $A_1, A_2, ..., A_t$.

To solve

$$\begin{aligned} \max \quad & \left(f_1^T M_k - c_k^T\right) x_k \\ \text{subject to} \quad & \\ & A_k x_k = b_k, k = 1, ..., t \\ & x_k \geq 0, k = 1, ..., t, \end{aligned} \tag{4.5.1}$$

we look at the sign of the last component of f_1. If it is positive, the $f_1^T M_k$ is a higher order convex function (the order is the largest power in M_k) and if the sign is negative, the $f_1^T M_k$ is a higher order concave function.

We can optimize both objective functions, but if we want to solve only the minimization problem, then we multiply the objective function by -1, and minimize the modified objective function.

Assume that the original objective function determined by $(c_1^T, c_2^T, ..., c_t^T)$ is piecewise linear, i.e., each component c_i^T is linear, $i = 1, ..., t$. The linear term c_T in the objective function doesn't change higher order convexity or concavity, also the order remains the same.

Consider the special case, where

$$M_1, M_2, ..., M_t$$

are matrices of dimensions

$$1 \times m_1, 1 \times m_2, ..., 1 \times m_t.$$

Here there is just one coupling constraint, which is the sum of the x-components is equal to 1.

In this case f_1 has a single component and the objective function in (**??**) is

$$\left(f_1 - c_{k1}\right) x_{k1} + ... + \left(f_1 - c_{kn_k}\right) x_{kn_k},$$

thus the function of the objective coefficients is

$$f_1 - c_{k1}, ..., f_1 - c_{kn_k},$$

which is linear.

Consider, in general, the problem:

$$\begin{aligned} \min(\max) \quad & l^T x \\ \text{subject to} \quad & \\ & Ax = b \\ & x \geq 0, \end{aligned} \tag{4.5.2}$$

where A is in the form

$$A = \begin{pmatrix} v_0^h & v_1^h & \dots & v_M^h \\ v_0^{h+1} & v_1^{h+1} & \dots & v_M^{h+1} \\ \dots & \dots & \dots & \dots \\ v_0^{h+k} & v_1^{h+k} & \dots & v_M^{h+k} \end{pmatrix}.$$

where $M \geq k$ and l is linear on the set $v_0, v_1, ..., v_M$. Preferably, $v_0, v_1, ..., v_M$ are consecutive integers, but not necessarily starting with 0, but starting with an arbitrary integer.

Note that problem (4.5.2) just a general formulation of a subproblem (**??**).

Problem (4.5.2) is totally positive in the sense described in [65], provided that l is non-decreasing and

$$l(v) = av_i + b, a \geq 0.$$

If $a < 0$, we take the negative of the objective function. Obviously, we can remove b and can take $a = 1$ to obtain the optimal basis.

The dual feasible basis structures are introduced in [65]:

	k even	k odd
min problem	$\{i, i+1, \dots, k, k+1\}$	$\{i, i+1, \dots, k, k+1, M\}$
max problem	$\{0, i, i+1, \dots, k, k+1, M\}$	$\{0, i, i+1, \dots, k, k+1\}$

Accordingly, we can carry out the optimization in (4.5.2) and in our problem (4.5.1).

Having the incoming vector, the outgoing vector can easily be determined from problem (4.4.5).

4.6 Numerical results

First, we benchmark our result against a version of discrete normal distribution for bounds on $P(X >= 1)$ and $E(f(X))$, where f is a discontinuous function with constant value up to a point and then higher order convex.

- The moments are generated by standard normal distribution.
- The objective function is

$$f(x) = \begin{cases} e^{0.1x} + e^{0.5x} & \text{if } x \geq -2 \\ 0 & \text{otherwise} \end{cases}$$

First, we are going to include conditional moments of the normal distribution. Naturally, we will use the conditional moments at -2 since it is the discontinuity in the objective function. Then, we have $E(X|X < -2) = -0.053991$ and $E(X^2|X < -2) = 0.130732$.

Note that another advantage of using small number of conditional moments (< 6 moments) is the possibility of not using the high precision toolbox.

2 conditional moments: LB: 2.1042 UB: 2.3157 (without unimodal)

2 conditional moments: LB: 2.1216 UB: 2.2329 (with unimodal)

Now, we incorporate the information on the third conditional moment $E(X^3|X < -2) = -0.323946$.

3 conditional moments: LB: 2.1341 UB: 2.1677 (without unimodal)

3 conditional moments: LB: 2.1360 UB: 2.1482 (with unimodal)

Chapter 5

Applications of the discrete moment problem. Bounding the prices of financial derivatives

5.1 Overview of published work

Probability bounds based on discrete moment problems have been applied to a large number of practical problems. Among these which use an LP formulation and higher order moments we mention

- Finding the reliability bounds for transportation systems (bounding the probability of the existence of the feasible flow). See, for example, [8].
- Bounding reliability of communication networks (the probability that two customers can be connected with all the way working path, or that all customers should be connected, that is, two-terminal reliability or all terminal reliability). See, for example, [70].
- Bounding the values of financial derivatives. See, for example, [68], [6].
- Bounding the probability in the probabilistic constrained model in stochastic programming. See, for example, [67].
- Applications in natural sciences. See, for example, [90], [21], [4].
- Applications in numerical analysis, probability and statistics. See , for example,[10], [29], [86].

In this section we present some applications in economics and finance.

One of the first papers in the area was published Ritchken (1985) [74, 75], where the future security prices are supposed to form a finite element discrete set: $s_1, \ldots, s_n$

with known probabilities $p_1, \ldots, p_n$, respectively. The lower and upper bounds for the European call option are provided by the optimum values of the minimization and maximization problems:

$$\begin{aligned} \min \ (\max) \ & \sum_{i=1}^{n} [s_i - X]_+ p_i \\ \text{subject to} \ & \sum_{i=1}^{n} p_i = 1, \\ & \sum_{i=1}^{n} s_i p_i = \mu, \\ & p \geq 0, \end{aligned} \tag{5.1.1}$$

where X is the strike price. In case of the put option we have $[X - s_i]_+$ in the objective function.

Ritchken [74, 75] shows that the lower bound is attained at a distribution, where for a consecutive pair of probabilities p_i, p_{i+1} we have that $p_i + p_{i+1} = 1$, $p_i \geq 0$, $p_{i+1} \geq 0$, which implies $p_i = 0$ for $j \notin \{i, i+1\}$ and the upper bound is attained at the distribution, where $p_1 + p_n = 1$, $p_1 \geq 0$, $p_n \geq 0$. The phenomenon that a consecutive pair of $p_1, ..., p_n$ forms an optimal basic solution was also observed in connection with hydraulic systems, see Fujiwara [24], Khang and Fujiwara [51]. It is also a special case of basic solutions in discrete moment problems as we can see it in Chapter 2.

Lo (1987) [55] looked at the case, where the asset price process $S(t), t \geq 0$ is arbitrary, where the random variables have finite variances. He formulated the optimization problem:

$$\max_{F \in \mathfrak{F}} \mathbb{E}_F([S(T) - X]_+ | S(t)), \tag{5.1.2}$$

where X is the striking price and $\mathfrak{F}$ is the collection of conditional probability distributions of $S(T)$, given $S(t)$ such that

$$\mathbb{E}(S(T)|S(t)) = \mu(t, T), \tag{5.1.3}$$

$$Var(S(T)|S(t)) = \mathbb{E}(S^2(T)|S(t)) - \mu^2(t, T) = \sigma^2(t, T).$$

The solution to problem (5.1.2) exists and the optimum value provides us with an upper bound on the conditional expectation, where we have the conditional moment information (5.1.3):

$$\mathbb{E}[(S(T)-X)|S(t)] \leq$$

$$\begin{cases} \mu(t,T)(\mu(t,T)-X)+\mu(t,T)\sigma^2(t,T), & \text{if } X \leq \frac{\mu^2(t,T)+\sigma^2(t,T)}{2\mu(t,T)} \\ \frac{1}{2}(\mu(t,T)-X+\sqrt{(X-\mu(t,T))^2+\sigma^2(t,T)}), & \text{if } X > \frac{\mu^2(t,T)+\sigma^2(t,T)}{2\mu(t,T)} \end{cases} \tag{5.1.4}$$

To obtain this discounted conditional expectation we multiply by $e^{-r(T-t)}$ on both sides in (5.1.4). In case of time homogeneity we have $\mu(t,T)=\mu(T-t)$, $\sigma^2(t,T)=\sigma^2(T-t)$, where $\mu,\sigma^2>0$ are constants.

The discounted bound in (5.1.4) is applied in Lo (1987) [55] to obtain upper bounds for the European call in case of two special stochastic processes: (a) the multiplicative Brownian motion process and (b) Merton's mixed jump-diffusion process, under risk neutrality condition.

Grundy (1991) [34] gave upper bound on the expected payoff $\mathbb{E}([S(T)-X]_+)$, where the expectation of the nth moment of $S(T)$ is known. If it is $\mathbb{E}(S^n(T))=S_t^n\psi$, then we have the relations

$$\mathbb{E}([S(T)-X]_+) \leq \begin{cases} \psi^{\frac{1}{n}}S(t)-X, & \text{if } X \leq S(t)\frac{n-1}{n}\psi^{\frac{1}{n}} \\ \psi\frac{S(t)}{n}(\frac{S(t)}{X}\frac{n-1}{n})^{n-1}, & \text{if } X > S(t)\frac{n-1}{n}\psi^{\frac{1}{n}} \end{cases} \tag{5.1.5}$$

Grundy [34] also proved that if the asset price follows the dynamics:

$$dS(t)=\alpha(t)S(t)dt+\sigma(S(t),t)dB(t),$$

where $B(t), t\geq 0$ is the standard Brownian motion and $\alpha(t),\sigma(y,t)$ are deterministic functions, $\alpha(t)\geq r(t), t\geq 0$, then the expected return on the call payoff in a decreasing function of the time. Here $r(t)$ is the time dependent intensity of the rate of return.

Zheng (1994) formulated the moment problem, with the first two moments known, where the expected call payoff is to be maximized. The problem is

$$\begin{aligned}
\max \qquad & \int_0^\infty [z-X]_+ dF(z) \\
\text{subject to} \quad & \int_0^\infty dF(z) = 1, \\
& \int_0^\infty z dF(z) = \mu, \\
& \int_0^\infty z^2 dF(z) = \mu + \sigma^2, \\
& F \text{ is a c.d.f.,}
\end{aligned} \tag{5.1.6}$$

where F is the distribution function of the asset price at expiration. It is unknown but known are its first two moments μ and $\mu^2 + \sigma^2$, where σ^2 is the variance. The bounds are taken from the collection of formulas known in connection with the moment problem (see, e.g., Karlin, Studden, [47], Krein, Nudelman, [58], Prékopa, [67]).

5.2 Finding moment bounds on the European call option

In this section we present the moment bounds on the European call option based on DMP.

Let us introduce the following notation:

t = current time;

T = maturity of the option;

$S(t)$ = price of the underlying asset now;

$S(T)$ = random future price of the underlying asset;

K = striking price;

r = rate of interest, assuming continuous compounding;

c = price of the option.

Under risk neutral valuation, then the option price is the following:

$$c = e^{-r(T-t)}E([S(T) - K]_+|S(t) = s). \tag{5.2.1}$$

The Black-Scholes formula gives the value of c for the case where $S(\tau)$, $\tau \geq 0$ has the form

$$S(\tau) = S(0)e^{\sigma Z(\tau)+\mu\tau}, \tag{5.2.2}$$

where $Z(\tau)$, τ is the standard Brownian motion process, that is,

- $Z(0) = 0$;
- the process has independent increments;
- $Z(\tau)$ has distribution $N(0, \tau)$ and $\sigma > 0$, μ are constants.

The process (5.2.2) is a multiplicative Brownian motion.

We drop the assumption that $S(\tau)$ is a multiplicative Brownian motion process, let $t = 0$, and for simplicity, assume that $S(T)$ has the form

$$S(T) = e^{\alpha Z+\beta},$$

where Z is a random variable with support $0, \Delta, 2\Delta, ..., n\Delta$ with $\Delta > 0$. We also assume that there exists h such that

$$e^{\alpha h\Delta+\beta} - K = 0. \tag{5.2.3}$$

Thus, the payoff is 0, if $Z \leq h\Delta$, and it is $e^{\alpha k\Delta} - K$, if $Z = k \geq h$.

We consider two different ways for bounding options by the use of the moment problem. The first method is the following. Based on the knowledge of the first m conditional moments of Z given that $e^{\alpha Z+\beta} > K$, we can obtain the lower and upper bounds on the option price c. Let designate those conditional moments as $v_1, ..., v_m$, and $v_0 = 1$, by definition.

Since

$$\begin{aligned} c &= e^{-rT}E\left[|S(T) - K|_+\right] \\ &= e^{-rT}E\left[e^{\alpha Z+\beta} - K|e^{\alpha Z+\beta} > K\right]P(e^{\alpha Z+\beta} > K), \end{aligned} \tag{5.2.4}$$

we use DMP to find the lower and upper bounds for the factors in (5.2.4), for c.

The bounds for $E\left[e^{\alpha Z+\beta} - K|e^{\alpha Z+\beta} > K\right]$ are, therefore, the optimal values for the following linear programs:

$$\min\ (\max) \sum_{i=h}^{n} \left(e^{\alpha i\Delta+\beta} - K\right) x_i$$

$$\text{subject to} \tag{5.2.5}$$

$$\sum_{i=h}^{n} (i\Delta)^k x_i = v_k, k = 0, 1, ..., m \tag{5.2.6}$$

$$x_i \geq 0, i = h, ..., n.$$

Note that the objective function has positive divided differences of all orders on the support set $h, ..., n$.

Therefore, there is no limit on the number of moments one can use solving for the bounds.

The lower and upper bounds on the option price c are given in the Table below for the case of $\alpha = 1$, $\beta = 0$, $\Delta = 0.15$, $K = 4.48$, $T = 20$, $r = 0.5$.

Table 5.1: Lower and upper bounds on an European option

Number of moments	LB	UB
2	0.098838092430328	1.6347640805735
3	0.252180335760311	1.1183358715399
4	0.352590436816222	0.830771645118
5	0.426106557290974	0.701620079617
6	0.470554589696063	0.647768606808
7	0.49141209253348	0.631292204563
8	0.497177792126569	0.626644284407
9	0.502329424759515	0.615894953434
10	0.512831125580792	0.609569285522
11	0.518921786607432	0.607588751856
12	0.521342727169581	0.607330103374
13	0.521451366024903	0.603042392533
14	0.526956225610042	0.598975217350
15	0.53079421002388	0.597182236940
16	0.533061752760335	0.596730832159
17	0.533654316831338	0.595314663257
19	0.538382274919909	0.590699009320
20	0.540259140026421	0.590306260034

The underlying stock price follows a uniform distribution on $0, \Delta, 2\Delta, ..., n\Delta$, where $n = 80$.

Chapter 6

Conclusions

We have presented some new applications and numerical procedures for the univariate power moment problem in a finite interval. Significant improvements in both of the lower and upper bounds in the continuous moment problems can be obtained if the shape constraints are introduced, in addition to the moment matching constraints in the LP of the DMP. Another way to improve on the bounds, making them more suitable for approximation of the expectation of a function of a random variable, is the use of DMP in the continuous moment problem.

We present a detailed description of this methodology under the condition that the function is piecewise higher order convex. Here the bounding polynomials are allowed to run above (below) the function in the minimization (maximization) problem on small intervals (deviation in the wrong direction) but this property serves to get closer the lower and upper bounds in a controlled manner.

The elegant and efficient algorithms for numerical integration, worked out for the discrete case, allow for the solution of both problem types, where the number of known moments can be very large. Linear programming theory is applied and special LP algorithms are developed to efficiently solve the problems. Illustrative numerical examples are presented, where it is shown that our methods can solve problems with a significant number of the given moments.

We also consider an application of the DMP for the solutions of the discrete conditional moment problems (the problems with the block structure). We demonstrate with an example how our method can effectively be applied in practice.

Finally, we introduce an application of the DMP to option pricing valuation. A small example demonstrates the case of European option price valuation given 20 moments with the error bound 10^{-2}.

References

[1] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions.* Dover, Mineola, NY, 1965.

[2] N.I. Akhiezer. *The classical moment problem.* Oliver and Boyd, Edinburgh, 1965.

[3] N.I. Akhiezer and M.G.Krein. *Some questions in the theory of moments, vol.2 of* Transitions of Mathematical monographs. American Mathematical Society, Providence, RI, 1968.

[4] A. Antolín, Zarzo A., and Angulo J.C. Upper and lower bounds on the radial electron density in atoms. *Physical review A*, 48(6):4149–4155, 1993.

[5] D.H. Bailey. High-precision computation: applications and challenges, keynote address. *21st IEEE International Symposium on Computer Arithmetic (Arith 21) in Austin, Texas in April 2013.*

[6] D. Bertsimas and I. Popescu. On the relation between option and stock prices: A convex optimization approach. *Operations Research*, 50(2):358–374, 2002.

[7] I. Bienaymé. Considérations a l'appui de la découverte de Laplace sur la loi de probabilité dans la méthode des moindres carrés. *Comptes Rendus Académie des Sciences, Paris*, 37:309–326, 1853.

[8] E. Boros and A. Prékopa. On the existence of a feasible flow in a stochastic transportation network. *Operations Research*, 39:119–129, 1991.

[9] E. Boros and A. Prékopa. Closed form two-sided bounds for probabilities that exactly r and at least r out of n events occur. *Generalized Convexity and Generalized Monotonicity, Lecture Notes in Economics and Mathematical Systems under N. Hadjisavvas, J.E. Martinez-Legaz, J-P. Penot, eds.*, 502:293–320, 2001.

[10] J. Bukszár. Hypermultitrees and sharp Bonferroni inequalities. *Math. Inequal. Appl.*, 6(4):727–743, 2003.

[11] T. Carleman. Sur le problème des moments. *Comptes Rendus*, 174:1680–1682, 1922.

[12] A. Charnes, W. Cooper, and K. Kortanek. Duality, Haar programs and finite sequence spaces. *Proc. Natl. Acad. Sci. USA*, 48:783–786, 1962.

[13] A. Charnes, W. Cooper, and K. Kortanek. On representation of semi-infinite programs which have no duality gaps. *Management Science*, 12:113–121, 1965.

[14] P. Chebyshev. Sur les valeurs limites des intégrales. *Journal de Mathématiques pures et appliquées*, 19:157–160, 1874.

[15] P. Chebyshev. Sur deux théorèmes relatifs aux probabilités. *Acta Mathematica*, 14:305–315, 1887.

[16] H. Cramér. Random variables and probability distributions. *Cambridge Tracts in Mathematics and Mathematical Physics*, 36, 1937.

[17] G. Dahlquist. On summation formulas due to Plana, Lindelöf and Abel, and related Gauss-Christoffel rules. *BIT*, 37(2):256–295, 1997.

[18] G. Dahlquist and Å. Björck. *Numerical methods.* Dover publications, New York, 1974.

[19] P.J. Davis and P. Rabinowitz. *Methods of numerical integration, 2nd ed.* Academic Press, Orlando, Florida, 1984.

[20] D. A. Dawson and D. Sankoff. An inequality for probabilities. *Proceedings of the American Mathematical Society*, 18(3):501–507, 1967.

[21] A. Dhurandhar and A.Dobra. Semi-analytical method for analyzing models and model selection measures based on moment analysis. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 3(2):1–51, 2009.

[22] H. Engels. *Numerical quadrature and cubature.* Computational Mathematics and Applications, Academic Press, New York, 1980.

[23] T.O. Espelid. Doubly adaptive quadrature routines based on Newton-Cotes rules. *BIT*, 43(2):319–337, 2003.

[24] O. Fujiwara. Two adjacent pipe diameters at the optimal solution in the water distribution network models. *Water Resources Research*, 23(8):1457–1460, 1987.

[25] W. Gautschi. On the construction of Gaussian quadrature rules from modified moments. *Math.Comp.*, 24:245–260, 1970.

[26] W. Gautschi. *A survey of Gauss-Christoffel quadrature formulae. In P.L. Butzer and F. Fehér. E.B. Christoffel: The influence of his work on mathematics and the physical sciences.* Birkhäuser, Basel, 1981.

[27] W. Gautschi. On the computation of generalized FermiDirac and BoseEinstein integrals. *Computer Physics Communications*, 74:233–238, 1993.

[28] W. Gautschi. Orthogonal polynomials: applications and computation. *Acta Numerica*, 5:45–119, 1996.

[29] A. Genz. Numerical computation of multivariate normal probabilities. *Journal of Computational and Graphical Statistics*, 1(2):141–149, 1992.

[30] G.H. Golub and J. Kautsky. Calculation of Gauss quadratures with multiple free and fixed knots. *Numer.Math.*, 41:147–163, 1983.

[31] G.H. Golub and G. Meurant. *Matrices, moments and quadrature with applications.* Princeton University Press, 2010.

[32] G.H. Golub and J.H. Welsch. Calculation of Gauss quadrature rules. *Math.Comp.*, 23:221–230, 1969.

[33] P. Gonnet. Increasing the reliability of adaptive quadrature using explicit interpolants. *ACM Transactions on Mathematical Software*, 37(3 (26)), 2010.

[34] B. D. Grundy. Option prices and the underlying asset's return distribution. *The Journal of Finance*, 46(3):1045–1070, 1991.

[35] A. Haar. Über lineare Ungleichungen. *Acta Scientarium Mathematicarium (Szeged)*, 2:1–14, 1926.

[36] H. Hamburger. Beiträge zur Konvergenztheorie der Stieltjesschen Kettenbrüche. *Mathematische Zeitschrift*, 4:186–222, 1919.

[37] H. Hamburger. Über eine Erweiterung des Stieltjesschen Momentenproblems. *Math. Ann.*, 81:235–319, 1920.

[38] H. Hamburger. Über eine Erweiterung des Stieltjesschen Momentenproblems. *Math. Ann.*, 82:120–164, 168–187, 1921.

[39] F. Hausdorff. Summationsmethoden und Momentfolgen. *Mathematische Zeitschrift*, 16:74–109, 280–299, 1921.

[40] F. Hausdorff. Momentprobleme für ein endliches Intervall. *Mathematische Zeitschrift*, 9:220–248, 1923.

[41] E.K. Haviland. On the momentum problem for distributions in more than one dimension. *American Journal of Mathematics*, 57:562–568, 1935.

[42] E.K. Haviland. On the momentum problem for distribution functions in more than one dimension ii. *American Journal of Mathematics*, 58:164–168, 1936.

[43] F.B. Hildebrand. *Introduction to Numerical Analysis.* Dover publications, New York, 1974.

[44] K. Isii. The extrema of probability by generalized moments. *Ann. Inst. Stat. Math.*, 12:119–133, 1960.

[45] K. Isii. On sharpness of Tchebysheff type inequalities. *Ann. Inst. Stat. Math.*, 14:185–197, 1963.

[46] S. Karlin and L.S. Shapley. *Geometry of moment spaces. Number 12 in Memoirs of the American Mathematical Society.* American Mathematical Society, Providence, RI, 1953.

[47] S. Karlin and W.J. Studden. *Tchebysheff systems: with applications in Analysis and Statistics, vol.XV.* Pure and Applied Mathematics, Interscience publishers, NY, 1966.

[48] J. Kautsky and S. Elhay. Calculation of the weights of interpolatory quadratures. *Numer.Math.*, 40:407–422, 1982.

[49] J. Kautsky and S. Elhay. Algorithm 655: Iqpack, fortran subroutines for the weights of interpolatory quadrature. *ACM Transactions on Mathematical Software*, 13(4):399–415, 1987.

[50] J.H.B. Kemperman. The general moment problem, a geometric approach. *Annals of Mathematical Statistics*, 39:93–122, 1968.

[51] D.B. Khang and O. Fujiwara. Optimal adjacent pipe diameters in water distribution networks with reliability constraints. *Water Resources Research*, 28(6):1503–1505, 1992.

[52] R. Kim and M. S. Lundstrom. Notes on FermiDirac integrals. *[online] Available: http://nanohub.org/resources/5475*, 2008.

[53] T. H. Kjeldsen. The early history of the moment problem. *Historia Mathematica*, 20:19–44, 1993.

[54] D.P. Laurie. Computation of Gauss-type quadrature formulas. *J. Comp. Appl. Math.*, 127:201–217, 2001.

[55] A. Lo. Semi-parametric upper bounds for option prices and expected payoffs. *Journal of Financial Economics*, 19(2):373–388, 1987.

[56] J. Lyness. Notes on the adaptive Simpson quadrature routine. *J. Assoc. Comput.Mach.*, 16:483–495, 1969.

[57] A. Markov. *On certain applications of algebraic continued fractions*. Doctoral Thesis, St.Petersburg, 1884.

[58] M.G.Krein and A.A.Nudelman. The Markov moment problem and extremal problems. *Library of Congress Cataloging in Publication Data*, 50, 1977.

[59] I.P. Mysovskih. On the construction of the cubature formulas with the smallest number of nodes. *Soviet Math. Dokl.*, 9:277–280, 1968.

[60] R. Nevanlinna. Asymptotische Entwicklungen beschränkter Funktionen und das Stieltjessche Momentenproblem. *Ann. Acad. Scient. Fennicae*, (A), 18(5):1–53, 1922.

[61] R. Piessens, E. de Doncker nd C.W. Überhuber, and D.K. Kahaner. *QUADPACK: A subroutine package for automatic integration.* Springer Verlag, 1983.

[62] A. Prékopa. Boole-Bonferoni inequalities and linear programming. *Operations Research*, 36, No.1:145–162, 1988.

[63] A. Prékopa. The discrete moment problem and linear programming. *Discrete Applied Mathematics*, 27:235–254, 1990.

[64] A. Prékopa. Sharp bounds on probabilities using linear programming. *Operations Research*, 38:227–239, 1990.

[65] A. Prékopa. *Totally Positive Linear Programming Problems, In: Functional Analysis, Optimization and Mathematical Economics. A Collection of Papers Dedicated to the Memory of L.V. Kantorovich (L.J. Leifman, ed.).* Oxford University Press, New York, Oxford, 1990.

[66] A. Prékopa. Inequalities on expectations based on the knowledge of multivariate moments. *In: Stochastic Inequalities (M. Shaked and Y.Z.Tong eds.), Institute of Mathematical Statistics, Lecture Notes - Monograph Series*, 22:309–331, 1992.

[67] A. Prékopa. *Stochastic Programming.* Kluwer Scientific Publishers, Boston, 1995.

[68] A. Prékopa. Discrete higher order convex functions. *Generalized Convexity and Generalized Monotonicity, Lecture Notes in Economics and Mathematical Systems under N. Hadjisavvas, J.E. Martinez-Legaz, J-P. Penot, eds.*, 502:293–320, 2001.

[69] A. Prékopa and G. Alexe. Dual methods for the numerical solution of the univariate power moment problem. *Rutcor Research Report*, 14, 2003.

[70] A. Prékopa, E. Boros, and K.W. Lih. The use of binomial moments for bounding network reliability. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 5:197–212, 1991.

[71] H. Richter. Parameterfreie Abschätzung und Realisierung von Erwartungswerten. *Blätter der Deutschen Gesellschaft für Versicherungsmathematik*, 3:147–161, 1957.

[72] F. Riesz. Sur les opérations fonctionelles linéaires. *Comptes Rendus Académie des Sciences, Paris*, 149:974–977, 1909.

[73] F. Riesz. Sur le problème des moments. *Arkiv for matematik, Astronomi och Fysik (in French)*, 17:1–52, 1923.

[74] P.H. Ritchken. On option pricing bounds. *The Journal of Finance*, 40(4):1219–1233, 1985.

[75] P.H. Ritchken and S. Kuo. Option bounds with finite revision opportunities. *The Journal of Finance*, 43(2):301–308, 1988.

[76] W.W. Rogosinsky. Non-negative linear functionals, moment problems, and extremum problems in polynomial spaces. *Studies in mathematical analysis and related topics*, 316(324):121–125, 1962.

[77] S.M. Samuels and W.J. Studden. *Bonferroni-type probability bounds as an application of the theory of Tchebycheff System. Probability, Statistics and Mathematics, Papers in Honor of Samuel Karlin.* Academic Press, 1989.

[78] J.A. Shohat and D.J. Tamarkin. *The problem of moments. Mathematical surveys.* American Mathematical Society, Providence, RI, 1946.

[79] M. Sobel and V. R. R. Uppuluri. On bonferroni-type inequalities of the same degree for the probability of unions and intersections. *Ann. Math. Statist.*, 43(5):1549–1558, 1972.

[80] T.J. Stieltjes. Recherches sur quelques séries semi-convergentes. phd thesis. *Gauthier-Villars*, 1886.

[81] T.J. Stieltjes. Recherches sur les fractions continues. *Annales de la Faculté des Sciences de Toulouse*, 8:1–122, 1894.

[82] T.J. Stieltjes. Recherches sur les fractions continues. *Annales de la Faculté des Sciences de Toulouse*, 9:5–47, 1895.

[83] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis.* Springer, 1992.

[84] A.H. Stroud. *Approximate calculation of multiple integrals.* Prentice-Hall, Englewood Cliffs, NJ, 1972.

[85] A.H. Stroud and D. Secrest. *Gaussian Quadrature Formulas.* Prentice-Hall, Englewood Cliffs, NJ, 1966.

[86] Tamás Szántai. Improved bounds and simulation procedures on the value of the multivariate normal probability distribution function. *Annals OR*, 100(1-4):85–101, 2000.

[87] V. Tchakaloff. Formules de cubatures mécaniques à coefficients non négatifs. *Bull. Sci. Math.*, 81(2):123–134, 1957.

[88] M. Tortorella. Closed Newton-Cotes quadrature rules for Stieltjes integrals and numerical convolution of life distributions. *SIAM J. Stat. Comput.*, 11:732–748, 1990.

[89] U. von Matt. *Gauss quadrature.* Springer-Verlag, Berlin, 4th ed., 2004.

[90] J.C. Wheeler and R.G.Gordon. Bounds for averages using moment constraints. *Mathematics in Science and engineering*, 71:99–128, 1970.