Available online at www.sciencedirect.com
ScienceDirect
Operations Research Letters 36 (2008) 205–210
Operations Research Letters
www.elsevier.com/locate/orl

Approximating integrals of multivariate exponentials: A moment approach

Dimitris Bertsimas[a], Xuan Vinh Doan[b,*], Jean Lasserre[c]

[a] Sloan School of Management, Operations Research Center, Massachusetts Institute of Technology, E40-147, Cambridge, MA 02139-4307, USA
[b] Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, USA
[c] LAAS, 7 Avenue Du Colonel Roche, 31077 Toulouse Cédex 4, France

Received 5 February 2007; accepted 4 July 2007
Available online 11 August 2007

Abstract

We propose a method to calculate lower and upper bounds of some exponential multivariate integrals using moment relaxations and show that they asymptotically converge to the value of the integrals when the moment degree increases. We report computational results for integrals involving the normal distribution and exponential order statistic probabilities.

Keywords: Moment relaxation; Multivariate exponentials; Integral approximation; Semidefinite programming

1. Introduction

Multivariate integrals arise in statistics, physics, engineering and finance applications among other areas. For example, these integrals are needed to calculate probabilities over compact sets for multivariate normal random variables. It is therefore important to compute or approximate multivariate integrals. Usual methods include Monte Carlo schemes (see [10] for details) and cubature formulae as shown in e.g. de la Harpe and Pache [3]. Genz [4] presents very good algorithms for rectangular probability computation of bivariate and trivariate normal distributions. However, there are still many open problems currently and research on general multivariate integrals is very much active due to its importance as well as its difficulties. For instance, most cubature formulas are restricted to special sets like boxes and simplices, and even in this particular context, determination of orthogonal polynomials used to construct a! cubature is not an easy task.

1.1. Contributions and paper outline

In this paper, we attempt to approximate a class of exponential integrals, which in particular can be useful to calculate probabilities of multivariate normal random variables on compact sets $\Omega \subset \mathbb{R}^n$ more general than hyper rectangles boxes or simplices. Specifically, our contribution and structure of the paper are as follows:

(1) In Section 2, we provide a general framework to calculate lower and upper bounds for a class of exponential integrals on $\Omega \subset \mathbb{R}^n$. These bounds are calculated by solving a hierarchy of specific semidefinite programming problems constructed from appropriate sequences of moments.
(2) In Section 3, we prove that the two monotone sequences of lower and upper bounds generated by these semidefinite programming problems will asymptotically converge to the real value of the integral. The proof is due to some results from the problem of moments.
(3) In Section 4, computational results are reported for order statistic probabilities of multivariate random variables such as Gumbel's exponentials and normal variables. These results show that the proposed method is indeed applicable for this class of exponential integrals.

2. Moment framework

We consider the following class of multivariate exponential integrals:

$$\rho = \int_{\Omega} g(\boldsymbol{x}) \mathrm{e}^{h(\boldsymbol{x})}\, \mathrm{d}\boldsymbol{x}, \tag{1}$$

* Corresponding author.
E-mail addresses: dbertsim@mit.edu (D. Bertsimas), vanxuan@mit.edu (X.V. Doan), lasserre@laas.fr (J. Lasserre).

0167-6377/$ - see front matter
doi:10.1016/j.orl.2007.07.002

where $\boldsymbol{x} \in \mathbb{R}^n$, $g, h \in \mathbb{R}[\boldsymbol{x}]$, the ring of real polynomials, and $\Omega \subset \mathbb{R}^n$ is a compact set defined as

$$\Omega = \{\boldsymbol{x} \in \mathbb{R}^n : b_1^l \leqslant x_1 \leqslant b_1^u, b_i^l(\boldsymbol{x}[i-1]) \leqslant x_i \leqslant b_i^u(\boldsymbol{x}[i-1]) \quad \forall i = 2, \ldots, n\}, \tag{2}$$

where $\boldsymbol{x}[i] \in \mathbb{R}^i$ is the vector of first i elements of $\boldsymbol{x}$ for all $i = 1, \ldots, n$, $b_i^l, b_i^u \in \mathbb{R}[\boldsymbol{x}[i-1]]$ for all $i = 2, \ldots, n$, and $b_1^l, b_1^u \in \mathbb{R}$.

For clarity of exposition, we will only describe the approach for simple integrals in two variables x and y on a box $[a, b] \times [c, d]$ in this paper. The multivariate case $n \geqslant 3$ essentially uses the same machinery with more complicated and tedious notation. Its detailed exposition can be found in Bertsimas et al. [1].

2.1. Integral formulation

Suppose that one wants to approximate

$$\rho = \int_a^b \int_c^d g(x, y) e^{h(x,y)} \, dy \, dx, \tag{3}$$

where g, h are bivariate polynomials and where $\Omega := [a, b] \times [c, d]$.

Consider the measure μ on $\mathbb{R}^2$ defined by

$$\mu(B) = \int_{\Omega \cap B} e^{h(x,y)} \, dy \, dx \quad \forall B \in \mathscr{B}(\mathbb{R}^2), \tag{4}$$

and its sequence of moments $\boldsymbol{z} = \{z(\alpha, \beta)\}$:

$$z(\alpha, \beta) = \int x^\alpha y^\beta \, d\mu(x, y) = \int_\Omega x^\alpha y^\beta e^{h(x,y)} \, dy \, dx \quad \forall (\alpha, \beta) \in \mathbb{N}^2. \tag{5}$$

Clearly, $\rho = \sum_{(\alpha,\beta)\in\mathbb{N}^2} g_{(\alpha,\beta)} z(\alpha, \beta) = \langle \boldsymbol{g}, \boldsymbol{z} \rangle$, where $\boldsymbol{g} = (g_{(\alpha,\beta)})$ and $g_{(\alpha,\beta)}$ is the coefficient of the monomial $x^\alpha y^\beta$. Therefore we can compute ρ once we have all necessary moments $z(\alpha, \beta)$.

Integration by parts yields

$$z(\alpha, \beta) = \frac{1}{\beta+1} \int_a^b x^\alpha [y^{\beta+1} e^{h(x,y)}]_{y=c}^{y=d} \, dx - \frac{1}{\beta+1} \int_a^b \int_c^d x^\alpha y^{\beta+1} \frac{\partial h(x, y)}{\partial y} e^{h(x,y)} \, dy \, dx.$$

If $\{h_{(\gamma,\delta)}\}$ is the finite sequence of coefficients of the polynomial $h(x, y)$ then

$$z(\alpha, \beta) = \frac{d^{\beta+1}}{\beta+1} \int_a^b x^\alpha e^{h(x,d)} \, dx - \frac{c^{\beta+1}}{\beta+1} \int_a^b x^\alpha e^{h(x,c)} \, dx - \sum_{(\gamma,\delta)\in\mathbb{N}^2} \frac{\delta h_{(\gamma,\delta)}}{\beta+1} z(\alpha+\gamma, \beta+\delta).$$

Define the following two measures $d\nu = e^{h(x,d)} \, dx$ and $d\xi := e^{h(x,c)} \, dx$ on $[a, b]$ with their corresponding sequences of moments:

$$v(\alpha) = \int x^\alpha \, d\nu(x) = \int_a^b x^\alpha e^{h(x,d)} \, dx;$$
$$w(\alpha) = \int x^\alpha \, d\xi(x) = \int_a^b x^\alpha e^{h(x,c)} \, dx, \tag{6}$$

for all $\alpha \in \mathbb{N}$. Then

$$z(\alpha, \beta) = \frac{d^{\beta+1}}{\beta+1} v(\alpha) - \frac{c^{\beta+1}}{\beta+1} w(\alpha) - \sum_{(\gamma,\delta)\in\mathbb{N}^2} \frac{\delta h_{(\gamma,\delta)}}{\beta+1} z(\alpha+\gamma, \beta+\delta). \tag{7}$$

Let $k(x) := h(x, d)$ and $l(x) := h(x, c)$. Clearly, $x \mapsto k(x)$ and $x \mapsto l(x)$ are univariate polynomials in x. Integration by parts for $v(\alpha)$ in (6) yields

$$v(\alpha) = \frac{1}{\alpha+1} [x^{\alpha+1} e^{k(x)}]_{x=a}^{x=b} - \frac{1}{\alpha+1} \int_a^b x^{\alpha+1} \frac{dk(x)}{dx} e^{k(x)} \, dx \quad \forall \alpha \in \mathbb{N}$$

or

$$v(\alpha) = \frac{b^{\alpha+1} e^{k(b)}}{\alpha+1} - \frac{a^{\alpha+1} e^{k(a)}}{\alpha+1} - \sum_{\varepsilon\in\mathbb{N}} \frac{\varepsilon k_\varepsilon}{\alpha+1} v(\alpha+\varepsilon) \quad \forall \alpha \in \mathbb{N}, \tag{8}$$

where $\{k_\varepsilon\}$ is the finite sequence of coefficients of the polynomial $x \mapsto k(x) = \Sigma_\varepsilon k_\varepsilon x^\varepsilon$ of degree k_x.

Similarly, we have

$$w(\alpha) = \frac{b^{\alpha+1} e^{l(b)}}{\alpha+1} - \frac{a^{\alpha+1} e^{l(a)}}{\alpha+1} - \sum_{\varepsilon\in\mathbb{N}} \frac{\varepsilon l_\varepsilon}{\alpha+1} w(\alpha+\varepsilon) \quad \forall \alpha \in \mathbb{N}, \tag{9}$$

where $\{l_\varepsilon\}$ is the finite sequence of coefficients of the polynomial $x \mapsto l(x) = \Sigma_\varepsilon l_\varepsilon x^\varepsilon$ of degree l_x.

In view of (8) and (9), all $v(\alpha)$ and $w(\alpha)$ are affine functions of $v_0, \ldots, v_{k_x-1}$, and $w_0, \ldots, w_{l_x-1}$, respectively. In the next section, we introduce the moment relaxation framework based on necessary conditions of moment sequences to calculate the given integral.

2.2. Moment relaxation

Consider the measure μ and its corresponding sequence of moments $\boldsymbol{z}$. For every $r \in \mathbb{N}$, the r-moment matrix associated with μ (or equivalently, with $\boldsymbol{z}$) $M_r(\mu) \equiv M_r(\boldsymbol{z})$ is a matrix of size $\binom{r+2}{r}$. Its rows and columns are indexed in the canonical basis $\{x^\alpha y^\beta\}$ of $\mathbb{R}[x, y]$, and

$$M_r(\boldsymbol{z})((\alpha, \beta), (\gamma, \delta)) := z(\alpha+\gamma, \beta+\delta), \quad \alpha+\beta, \gamma+\delta \leqslant r. \tag{10}$$

Similarly, given $\theta \in \mathbb{R}[x, y]$, the localizing matrix $M_r(\boldsymbol{\theta z})$ associated with $\boldsymbol{z}$ and θ is defined by

$$M_r(\boldsymbol{\theta z})((\alpha, \beta), (\gamma, \delta)) := \sum_{(\varepsilon,\zeta)\in\mathbb{N}^2} \theta_{(\varepsilon,\zeta)} z(\alpha+\gamma+\varepsilon, \beta+\delta+\zeta), \quad \alpha+\beta, \gamma+\delta \leqslant r, \tag{11}$$

where $\boldsymbol{\theta} = \{\theta_{(\alpha,\beta)}\}$ is the vector of coefficients of θ in the canonical basis $\{x^\alpha y^\beta\}$.

If we define the matrix $M_r^{(\varepsilon,\zeta)}(\boldsymbol{z})$ for all $(\varepsilon, \zeta) \in \mathbb{N}^2$ with elements

$$M_r^{(\varepsilon,\zeta)}(\boldsymbol{z})((\alpha, \beta), (\gamma, \delta)) = z(\alpha+\gamma+\varepsilon, \beta+\delta+\zeta), \quad \alpha+\beta, \gamma+\delta \leqslant r,$$

then the localizing matrix can be expressed as $M_r(\boldsymbol{\theta z}) = \sum_{(\varepsilon,\zeta)\in\mathbb{N}^2}\theta_{(\varepsilon,\zeta)}M_r^{(\varepsilon,\zeta)}(\boldsymbol{z})$.

Note that for every polynomial $f \in \mathbb{R}[x, y]$ of degree at most r with its vector of coefficients denoted by $\boldsymbol{f}=\{f_{(\alpha,\beta)}\}$, we have

$$\langle \boldsymbol{f}, M_r(\boldsymbol{z})\boldsymbol{f}\rangle = \int f^2\,\mathrm{d}\mu, \quad \langle \boldsymbol{f}, M_r(\boldsymbol{\theta z})\boldsymbol{f}\rangle = \int \theta f^2\,\mathrm{d}\mu. \tag{12}$$

Then necessarily, $M_r(\boldsymbol{z})\succcurlyeq\mathbf{0}$ and $M_r(\boldsymbol{\theta z})\succcurlyeq\mathbf{0}$ whenever μ has its support contained in the level set $\{(x, y) \in \mathbb{R}^2 : \theta(x, y)\geqslant 0\}$. If the sequence of moments is restricted to those moments used to construct the moment matrix $M_r(\boldsymbol{z})$ (up to moments of degree $2r$), then the second necessary condition is reduced to $M_{r-\lceil d/2\rceil}(\boldsymbol{\theta z})\succcurlyeq\mathbf{0}$, where d is the degree of the polynomial θ. For more details on moment matrices, local matrices, and these necessary conditions, refer to Lasserre [7], Laurent [9] and references therein.

Define $\theta_1(x, y) := (b - x)(x - a)$ and $\theta_2(x, y) := (d - y)(y - c)$. Clearly, the support of μ is the semialgebraic set $\Omega=\{(x, y) \in \mathbb{R}^2 : \theta_i(x, y)\geqslant 0, \quad i = 1, 2\}$. As θ_1, θ_2 are both quadratic, the necessary conditions for moment and localizing matrices associated with μ read

$$M_r(\boldsymbol{z})\succcurlyeq 0, \quad M_{r-1}(\theta_i\boldsymbol{z})\succcurlyeq 0, \ \ i = 1, 2. \tag{13}$$

For the one-dimensional sequences $\boldsymbol{v}$ and $\boldsymbol{w}$ associated with v and ξ, one has obvious analogue definitions of moment and localizing matrices. Both measures v and ξ are supported on the set $\Psi=\{x \in \mathbb{R} : \theta_3(x)\geqslant 0\}$, with $\theta_3(x) := (b - x)(x - a)$, and so, analogues of (13) can be derived for $\boldsymbol{v}$ and $\boldsymbol{w}$.

Combining these necessary conditions and the linear relations for $\boldsymbol{z}$, $\boldsymbol{v}$, and $\boldsymbol{w}$ in (7)–(9), one obtains a lower bound for ρ by solving the following semidefinite programming problem $\mathscr{P}_r^l$:

$$\begin{aligned}
Z(\mathscr{P}_r^l) = \inf_{\boldsymbol{z},\boldsymbol{v},\boldsymbol{w}} \quad & \langle \boldsymbol{g}, \boldsymbol{z}\rangle \\
\text{s.t.} \quad & M_r(\boldsymbol{z})\succcurlyeq 0, M_r(\theta_i\boldsymbol{z})\succcurlyeq 0, \quad i = 1, 2, && \text{(14a)}\\
& M_r(\boldsymbol{v})\succcurlyeq 0, M_r(\theta_3\boldsymbol{v})\succcurlyeq 0, && \text{(14b)}\\
& M_r(\boldsymbol{w})\succcurlyeq 0, M_r(\theta_3\boldsymbol{w})\succcurlyeq 0, && \text{(14c)}\\
& z(\alpha, \beta) = \frac{d^{\beta+1}}{\beta+1}v(\alpha) - \frac{c^{\beta+1}}{\beta+1}w(\alpha) \\
& \qquad - \sum_{(\gamma,\delta)\in\mathbb{N}^2}\frac{\delta h_{(\gamma,\delta)}}{\beta+1}z(\alpha+\gamma, \beta+\delta), \\
& \forall(\alpha, \beta) \in \mathbb{N}^2 : \alpha+\gamma+\beta+\delta\leqslant 2r \\
& \forall(\gamma, \delta) : h_{(\gamma,\delta)} \neq 0, && \text{(14d)}\\
& v(\alpha) = \frac{b^{\alpha+1}\mathrm{e}^{k(b)}}{\alpha+1} - \frac{a^{\alpha+1}\mathrm{e}^{k(a)}}{\alpha+1} \\
& \qquad - \sum_{\varepsilon\in\mathbb{N}}\frac{\varepsilon k_\varepsilon}{\alpha+1}v(\alpha+\varepsilon), \\
& \forall\alpha \in \mathbb{N} : \alpha+\varepsilon\leqslant 2r \quad \forall\varepsilon : k_\varepsilon \neq 0, && \text{(14e)}\\
& w(\alpha) = \frac{b^{\alpha+1}\mathrm{e}^{l(b)}}{\alpha+1} - \frac{a^{\alpha+1}\mathrm{e}^{l(a)}}{\alpha+1} \\
& \qquad - \sum_{\varepsilon\in\mathbb{N}}\frac{\varepsilon l_\varepsilon}{\alpha+1}w(\alpha+\varepsilon), \\
& \forall\alpha \in \mathbb{N} : \alpha+\varepsilon\leqslant 2r \quad \forall\varepsilon : l_\varepsilon \neq 0. && \text{(14f)}
\end{aligned}$$

Similarly, an upper bound of ρ is obtained by solving the problem $\mathscr{P}_r^u$ with same feasible set as in $\mathscr{P}_r^l$ but with maximization objective $Z(\mathscr{P}_r^u) = \sup_{\boldsymbol{z},\boldsymbol{v},\boldsymbol{w}}\langle \boldsymbol{g}, \boldsymbol{z}\rangle$ instead. Clearly, $Z(\mathscr{P}_r^l)\leqslant\rho\leqslant Z(\mathscr{P}_r^u)$ and we next prove that $Z(\mathscr{P}_r^l) \uparrow \rho$ and $Z(\mathscr{P}_r^u) \downarrow \rho$ as $r \to \infty$.

3. Convergence

In order to prove the convergence of $Z(\mathscr{P}_r^l)$ and $Z(\mathscr{P}_r^u)$, we need to prove that the linear relations in (7)–(9) define moment sequences for all measures μ, v, and ξ. We start with one-dimensional moment sequences by the following lemma.

Lemma 1. *Let $\boldsymbol{v}$ be the moment sequence of some Borel measure v' on Ψ and assume that $\boldsymbol{v}$ satisfies* (8). *Then $v' = v$.*

Proof. According to (8),

$$v(\alpha) = \int x^\alpha\,\mathrm{d}v'(x) = \left[\frac{x^{\alpha+1}}{\alpha+1}\mathrm{e}^{k(x)}\right]_{x=a}^{x=b} - \int \frac{x^{\alpha+1}}{\alpha+1}k'(x)\,\mathrm{d}v'(x) \quad \forall\alpha \in \mathbb{N}, \tag{15}$$

but we also have

$$\int x^\alpha\mathrm{d}v(x) = \left[\frac{x^{\alpha+1}}{\alpha+1}\mathrm{e}^{k(x)}\right]_{x=a}^{x=b} - \int \frac{x^{\alpha+1}}{\alpha+1}k'(x)\,\mathrm{d}v(x) \quad \forall\alpha \in \mathbb{N}. \tag{16}$$

Consider the *signed* measure $\phi := v' - v$ on Ψ. From (15) and (16)

$$\int x^\alpha\mathrm{d}\phi(x) = -\int \frac{x^{\alpha+1}}{\alpha+1}k'(x)\,\mathrm{d}\phi(x) \quad \forall\alpha \in \mathbb{N}. \tag{17}$$

Let $x \mapsto p(x) := \sum_{j=1}^d f_j x^j$ so that $p'(x) = \sum_{j=0}^{d-1} f_{j+1}(j + 1)x^j$ for all x. From Eq. (17), one obtains

$$\int [p'(x) + p(x)k'(x)]\,\mathrm{d}\phi(x) = 0. \tag{18}$$

We now prove that (18) also holds for all continuous function $f = xg$, where g is continuously differentiable on Ψ. Recall that polynomials are dense in the space of continuously differentiable functions on Ψ under the sup-norm $\max\{\sup_{x\in\Psi}|f(x)|, \sup_{x\in\Psi}|f'(x)|\}$ (see e.g. [2,5]). Therefore, for every $\varepsilon>0$, there exist $p_\varepsilon \in \mathbb{R}[x]$ such that $\sup_{x\in\Psi}|g(x) - p_\varepsilon(x)|\leqslant\varepsilon$ and $\sup_{x\in\Psi}|g'(x) - p'_\varepsilon(x)|\leqslant\varepsilon$ simultaneously. As (18) is true for the polynomial $p = xp_\varepsilon$,

$$\begin{aligned}
&\int [f'(x) + f(x)k'(x)]\,\mathrm{d}\phi(x) \\
&\quad = \int [x(g(x) - p_\varepsilon(x))]'\,\mathrm{d}\phi(x) \\
&\qquad + \int x[g(x) - p_\varepsilon(x)]k'(x)\,\mathrm{d}\phi(x).
\end{aligned} \tag{19}$$

Using $[x(g(x) - p_\varepsilon(x))]' = (g(x) - p_\varepsilon(x)) + x(g'(x) - p'_\varepsilon(x))$ yields

$$|[x(g(x) - p_\varepsilon(x))]'|\leqslant \left(1 + \sup_{x\in\Psi}|x|\right)\varepsilon \quad \forall x \in \Psi,$$

and similarly,

$$|x[g(x) - p_\varepsilon(x)]k'(x)| \leqslant \sup_{x\in\Psi} |xk'(x)|\varepsilon \quad \forall x \in \Psi.$$

Therefore

$$\left|\int [f'(x) + f(x)k'(x)]\,\mathrm{d}\phi(x)\right| \leqslant \varepsilon\left(1 + \sup_{x\in\Psi}|x| + \sup_{x\in\Psi} |xk'(x)|\right)\int \mathrm{d}|\phi|.$$

As $M = (1+\sup_{x\in\Psi}|x| + \sup_{x\in\Psi}|xk'(x)|)\int \mathrm{d}|\phi|$ is finite and (19) holds for all $\varepsilon > 0$, one obtains

$$\int [f'(x) + f(x)k'(x)]\,\mathrm{d}\phi(x) = 0 \quad \forall f := xg, \quad g \in \mathbb{R}[x]. \tag{20}$$

Next, for an arbitrary polynomial $x \mapsto g(x) = \sum_{j=0}^{d} g_j x^j$, let $G(x) := \sum_{j=0}^{d}(g_j/(j+1))x^{j+1}$ so that $G' = g$, and let $f := G\mathrm{e}^{-k(x)}$. Observe that $f(x)/x$ is continuously and $f'(x) = g(x)\mathrm{e}^{-k(x)} - f(x)k'(x)$ for all x. Using (20) yields

$$\int g(x)\mathrm{e}^{-k(x)}\,\mathrm{d}\phi(x) = 0 \quad \forall g \in \mathbb{R}[x]. \tag{21}$$

If we let $\mathrm{d}\phi' := \mathrm{e}^{-k(x)}\,\mathrm{d}\phi$ then $\int f\,\mathrm{d}\phi' = 0$ for all continuous function f on Ψ because polynomials are dense in the space of continuous functions on Ψ. This in turn implies that ϕ' is the zero measure. In addition, as $\mathrm{e}^{-k(x)} > 0$ for all $x \in \mathbb{R}$, ϕ is also the zero measure, and so $v' = v$, the desired result. $\square$

Using Lemma 1, we can now prove similar results for the main moment sequences in $\mathbb{R}^2$.

Lemma 2. *Let* $\mathbf{v}$ *and* $\mathbf{w}$ *be the moment sequences of some Borel measures* v' *and* ξ' *on* Ψ *and assume that* $\mathbf{v}$ *and* $\mathbf{w}$ *satisfy* (8) *and* (9). *Let* $\mathbf{z}$ *be the moment sequence of some Borel measure* μ' *on* Ω *and assume that* $\mathbf{z}$ *satisfies* (7). *Then* $\mu' = \mu$.

Proof. By Lemma 1, $v' = v$ and $\xi' = \xi$. Then, according to (7),

$$\int x^\alpha y^\beta\,\mathrm{d}\mu'(x,y) = \frac{d^{\beta+1}}{\beta+1}v(\alpha) - \frac{c^{\beta+1}}{\beta+1}w(\alpha) - \int \frac{x^\alpha y^{\beta+1}}{\beta+1}\frac{\partial h(x,y)}{\partial y}\,\mathrm{d}\mu'(x,y). \tag{22}$$

Similarly,

$$\int x^\alpha y^\beta\,\mathrm{d}\mu(x,y) = \frac{d^{\beta+1}}{\beta+1}v(\alpha) - \frac{c^{\beta+1}}{\beta+1}w(\alpha) - \int \frac{x^\alpha y^{\beta+1}}{\beta+1}\frac{\partial h(x,y)}{\partial y}\,\mathrm{d}\mu(x,y). \tag{23}$$

Then if we consider the *signed* measure $\phi := \mu' - \mu$ on Ω, one obtains

$$\int x^\alpha y^\beta\,\mathrm{d}\phi(x,y) = -\int \frac{x^\alpha y^{\beta+1}}{\beta+1}\frac{\partial h(x,y)}{\partial y}\,\mathrm{d}\phi(x,y). \tag{24}$$

Using similar arguments as in the proof of Lemma 1, one obtains

$$\int \left[\frac{\partial f(x,y)}{\partial y} + f(x,y)\frac{\partial h(x,y)}{\partial y}\right]\mathrm{d}\phi(x,y) = 0 \tag{25}$$

for all functions $f = yg$, provided that g and $\partial g(x,y)/\partial y$ are continuous.

For every polynomial $g \in \mathbb{R}[x,y]$ let $G := yP$, where $P \in \mathbb{R}[x,y]$ and $\partial G(x,y)/\partial y = g(x,y)$ for all x, y. Then with $f := G\mathrm{e}^{-h(x,y)}$, one obtains

$$\int g(x,y)\mathrm{e}^{-h(x,y)}\,\mathrm{d}\phi(x,y) = 0 \quad \forall g \in \mathbb{R}[x,y]. \tag{26}$$

Again using similar arguments as in the proof of Lemma 1, one obtains the desired result $\mu' = \mu$. $\square$

With Lemmas 1 and 2, we can now prove the following convergence theorem:

Theorem 1. *Consider the semidefinite programming problems* $\mathscr{P}_r^l$ *and* $\mathscr{P}_r^u$ *defined in* (14a). *Then*

(i) $Z(\mathscr{P}_r^l)$ *and* $Z(\mathscr{P}_r^u)$ *are finite and in addition, both* $\mathscr{P}_r^l$ *and* $\mathscr{P}_r^u$ *are solvable for* r *large enough.*
(ii) *As* $r \to \infty$, $Z(\mathscr{P}_r^l) \uparrow \rho$ *and* $Z(\mathscr{P}_r^u) \downarrow \rho$.

Proof. Clearly, the collection of truncated sequences of moments of μ, v, and ξ forms a feasible solution for both problems $\mathscr{P}_r^l$ and $\mathscr{P}_r^u$. Moreover,

$$|z(\alpha,\beta)| \leqslant \int_\Omega |x^\alpha y^\beta \mathrm{e}^{h(x,y)}|\,\mathrm{d}x\,\mathrm{d}y \leqslant \sup_{(x,y)\in\Omega} |x^\alpha y^\beta \mathrm{e}^{h(x,y)}|\mathrm{vol}(\Omega).$$

Recall that $\Omega \subset \mathbb{R}^2$ is compact. Therefore $u(\alpha,\beta) := \sup_{(x,y)\in\Omega}|x^\alpha y^\beta \mathrm{e}^{h(x,y)}| = \max_{(x,y)\in\Omega}|x^\alpha y^\beta \mathrm{e}^{h(x,y)}|$ is finite. Similarly, upper bounds $u_1(\alpha)$ and $u_2(\alpha)$ are obtained for $|v(\alpha)|$ and $|w(\alpha)|$, respectively. Now consider $\mathscr{P}_r^l$ and $\mathscr{P}_r^u$ with the additional bound constraints $|z(\alpha,\beta)| \leqslant u(\alpha,\beta)$, $|u(\alpha)| \leqslant u_1(\alpha)$, and $|w(\alpha)| \leqslant u_2(\alpha)$ for all α, β.

(i) The feasible sets of these two modified problems are bounded and closed. The objective functions are linear and both problems are feasible. Therefore, they are both solvable and their optimal values are finite.
(ii) Let $\{\mathbf{z}_r, \mathbf{v}_r, \mathbf{w}_r\}$ be an optimal solution of $\mathscr{P}_r^l$ (with additional bound constraints) and complete these truncated sequences with zeros to make them become infinite sequences. As $|\mathbf{z}_r(\alpha,\beta)|$, $|\mathbf{v}_r(\alpha)|$ and $|\mathbf{w}_r(\alpha)|$are bounded uniformly in r, by a standard diagonal argument, there is a subsequence $\{r_m\}$ and infinite sequences $\{\mathbf{z}^*, \mathbf{v}^*, \mathbf{w}^*\}$ such that the *pointwise* convergences $\mathbf{z}_{r_m}(\alpha,\beta) \to \mathbf{z}^*(\alpha,\beta)$, $\mathbf{v}_{r_m}(\alpha) \to \mathbf{v}^*(\alpha)$ and $\mathbf{w}_{r_m}(\alpha) \to \mathbf{w}^*(\alpha)$ hold. This in turn implies

$$M_s(\mathbf{z}^*) \succcurlyeq 0, \quad M_s(\theta_i \mathbf{z}^*) \succcurlyeq 0, \quad i = 1, 2; \quad s = 0, 1, \ldots.$$

Similar conditions hold for $\mathbf{v}^*$ and $\mathbf{w}^*$. As Ω and Ψ are compact, by Putinar [11], there exist measures μ' on Ω

and ν' and ξ' on Ψ, such that $\mathbf{z}^*$, $\mathbf{v}^*$, and $\mathbf{w}^*$ are their respective moment sequences. In addition, by pointwise convergence, $\mathbf{z}^*$, $\mathbf{v}^*$, and $\mathbf{w}^*$ satisfy the linear relations (7), (8), and (9). Therefore , by Lemmas 1 and 2, $\mu'=\mu$, $\nu'=\nu$, and $\xi'=\xi$. And so,

$$\lim_{m\to\infty}\langle \mathbf{g}, \mathbf{z}_{r_m}\rangle = \langle \mathbf{g}, \mathbf{z}^*\rangle = \rho.$$

By definition of truncated moment and localizing matrices, every feasible solution of $\mathscr{P}^l_{r+1}$ generates a feasible solution of $\mathscr{P}^l_r$ with same value. Hence $\langle \mathbf{g}, \mathbf{z}_r\rangle \leqslant \langle \mathbf{g}, \mathbf{z}_{r+1}\rangle$ whenever $r \geqslant \deg(g)$, and so $\langle \mathbf{g}, \mathbf{z}_r\rangle \uparrow \rho$. Similar arguments can be applied for the modified problem $\mathscr{P}^u_r$.

So far, the results are obtained for problems $\mathscr{P}^l_r$ and $\mathscr{P}^u_r$ with additional bound constraints. In fact it is enough to use the bounds $u(0,0)$, $u_1(0)$ and $u_2(0)$, along with a more subtle argument similar to the one used in the proof of Theorem 2 in Lasserre [8]. In other words, with the only additional bounds $z(0,0)\leqslant u(0,0)$, $v(0)\leqslant u_1(0)$ and $w(0)\leqslant u_2(0)$, $\mathscr{P}^l_r$ and $\mathscr{P}^u_r$ are solvable when r is large enough and $Z(\mathscr{P}^l_r)\uparrow\rho$ as well as $Z(\mathscr{P}^u_r)\downarrow\rho$. $\square$

4. Computational results

In this paper, and for clarity of exposition, formulations and convergence proofs have been provided for simple integrals in two variables over a box $\Omega\subset\mathbb{R}^2$. As already mentioned, results for the more general integrals defined in (1) can be found in Bertsimas et al. [1]. With this approach, we can approximate not only integrals on hyper-rectangles but also many other like e.g. order statistic integrals over the set $\Omega=\{\mathbf{x}\in\mathbb{R}^n : a\leqslant x_1\leqslant\cdots\leqslant x_n\leqslant b\}$. To illustrate our moment approach, we have computed order statistic probabilities of Gumbel's bivariate exponential distribution and bivariate and trivariate normal distributions. The required semidefinite programming problems are coded in Matlab and solved using SeDuMi package [12]. All computations are done under a Linux environment on a Pentium IV 2.40 GHz with 1 GB RAM.

As shown in Kotz et al. [6], the density function of Gumbel's bivariate exponential distribution is

$$f(x,y)=[(1+\theta x)(1+\theta y)-\theta]\mathrm{e}^{-x-y-\theta xy}, \quad x,y\geqslant 0, \quad (27)$$

where $0\leqslant\theta\leqslant 1$. To show how the moment order r affects integral results, we choose an arbitrary θ in [0, 1] and compute the probability over the set $\Omega=\{0\leqslant x\leqslant y\leqslant 1\}$ using different r. The results reported in Table 1 is for $\theta=0.5$. The value $\bar{\rho}=\frac{1}{2}[Z(\mathscr{P}^u_r)+Z(\mathscr{P}^l_r)]$ is approximately 0.215448 while the error $\Delta\rho=\frac{1}{2}\left[Z(\mathscr{P}^u_r)-Z(\mathscr{P}^l_r)\right]$ decreases to 10^{-9} when $r=5$. If a very high accuracy is not needed, we can approximate this integral with $r=3$ or 4 in much less time as reported in Table 1 (where the computational time is the total time for solve both $\mathscr{P}^u_r$ and $\mathscr{P}^l_r$).

Table 1
Computational results for a Gumbel's bivariate order statistic probability

Moment degree	$r=2$	$r=3$	$r=4$	$r=5$
$\frac{1}{2}[Z(\mathscr{P}^u_r) - Z(\mathscr{P}^l_r)]$	6.5288E − 03	1.3016E − 04	1.4933E − 06	2.8974E − 09
Time (s)	0.38	0.59	1.28	6.23

The second distribution family that we consider is the normal family. The density function of a multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}=\mathbf{A}\mathbf{A}'$ is

$$f(\mathbf{x})=\frac{1}{(2\pi)^{n/2}\det(\boldsymbol{\Sigma})^{1/2}}\mathrm{e}^{-(1/2)(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}. \quad (28)$$

We have generated $\boldsymbol{\mu}$ and $\mathbf{A}$ with random elements in $[-1,1]$. Order statistic integrals of these distributions are then computed with variables restricted in $[-1,1]$. The error tolerance is set to be 5×10^{-5} (4-digit accuracy) and the moment order r is increased until this accuracy is achieved. For $n=2$ (bivariate distribution), we have computed integrals for 100 randomly generated distributions. The average moment order is $\bar{r}=3.64$, which means that in most instances we only need to use $r=3$ or 4. Similar to Gumbel's exponential distribution case, the average computational time is around 1 s with $r=4$ for the bivariate normal distributions. For $n=3$ (trivariate distributions), we again can solve the semidefinite programming problems with $r=4$ or 5 to achieve the required accuracy. However, in this case the computational time is large (around 250 and 4000 s for $r=4$ and 5, respectively). Clearly, in comparing with results obtained with algorithms specially designed for rectangular normal probabilities by Genz [4], our method for trivariate normal distributions is not competitive. However, recall that our framework is designed to accommodate a more general class of exponential integrals with a novel approach which permits to compute other types of probabilities such as order statistic probabilities. In addition, the framework we proposed could be further developed to calculate integrations not only with polynomials but also rational functions, which have been recently analyzed by Lasserre [8]. Computationally, one may further analyze the structure of semidefinite programming problems $\mathscr{P}^u_r$ and $\mathscr{P}^l_r$ to reduce the computational time. For example, using linear relations presented in Section 2 to eliminate some moment variables could reduce substantially the problem size in terms of number of variables.

Acknowledgments

We would like to thank the associate editor and two anonymous referees for their comments on the earlier version of this paper.

References

[1] D. Bertsimas, X.V. Doan, J. Lasserre. Multivariate exponential integral approximations: a moment approach, Technical Report, Operations Research Center, MIT, January 2006.

[2] C. Coatmélec, Approximation et interpolation des fonctions différentiables de plusieurs variables, Ann. Sci. É.N.S. 3^e Ser. 83 (4) (1966) 271–341.

[3] P. de la Harpe, C. Pache. Cubature formulas, geometrical designs, reproducing kernels, and Markov operators, in: International Conference on Group Theory, Gaeta, June 2003.

[4] A. Genz, Numerical computation of rectangular bivariate and trivariate normal and t probabilities, Stat. Comput. 14 (2004) 151–160.
[5] M.W. Hirsch, Differential Topology, Springer, New York, 1976.
[6] S. Kotz, N. Balakrishnan, N.M. Johnson, Continuous Multivariate Distribution, Models and Applications, vol. 1, second ed., Wiley, New York, 2000.
[7] J.B. Lasserre, Global optimization and the problem of moments, SIAM J. Optim. 11 (3) (2001) 796–817.
[8] J.B. Lasserre, The K-moment problem with densities, Math. Programming, 2007, to appear, ⟨http://dx.doi.org/10.1007/s10107-007-0118-4⟩.
[9] M. Laurent, Moment matrices and optimization over polynomials—a survey on selected topics. Preprint, September 2005.
[10] H. Niederreiter. Random number generator and quasi-Monte Carlo method, in: SIAM CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 63, 1992.
[11] M. Putinar, Positive polynomials on compact semi-algebraic sets, Indiana Univ. Math. J. 42 (1993) 969–984.
[12] J.F. Sturm, Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones, Optim. Methods Software 11–12 (1999) 625–653.