Some applications of moments and SDP-relaxations

J.B. Lasserre

LAAS-CNRS,

Semidefinite Programming and Applications,

MSRI, Berkeley, October 2002

- global optimization (with polynomials)
- systems of polynomial equations
- bounds on measures with moment conditions
- invariant measures

Consider the **global optimization** problem

$$\mathbb{P} \mapsto p^* \; := \; \min\{g_0(x) \,|\, g_i(x) \; \geq \; 0, \; i = 1, \ldots m\},$$

where $g_i(x) : \mathbf{R}^n \to \mathbf{R}$ are all real-valued polynomials. Let

$$\mathbb{K} \; := \; \{x \in \mathbf{R}^n \,|\, g_i(x) \geq 0, \; i = 1, \ldots m\},$$

be the feasible set. Let

$$1, x_1, x_2, \ldots x_n, x_1^2, x_1 x_2, \ldots, x_n^r,$$

be a basis (of dimension $s(r)$) of the vector space of real-valued polynomials of degree at most r, and in this basis write

$$p(x) \; = \; \sum_{\alpha \leq r} p_\alpha \, x^\alpha \; = \; p_\alpha \left[x_1^{\alpha_1} \ldots x_n^{\alpha_n}\right],$$

with $\alpha = \sum_{i=1}^n \alpha_i$, and $p \in \mathbf{R}^{s(r)}$ its vector of coefficients.

The univariate case

In this case one considers

$$\mathbb{P} \rightarrow p^* \;=\; \min_{x \in \mathbb{K}} g_0(x), \;\text{ with } \begin{cases} \mathbb{K} \;\equiv\; \mathbf{R} \\ \mathbb{K} \;\equiv\; \mathbf{R}_+ \\ \mathbb{K} \;\equiv\; [a,b] \end{cases}$$

With $\mathbb{K} \equiv \mathbf{R}$, Shor (1987) was the first to show that $\mathbb{P}$ is a convex problem. Later Nesterov (1997) proposed an LMI formulation for the three cases.

The multivariate case is NP-hard in general.

Recent approaches (Lasserre, Nesterov, Parrilo,...) use results from

- (real) algebraic geometry (**positive polynomials**)
- functional analysis (**moments**),
- convex analysis (**SDP**)

Two dual points of view

p^* global minimum $\Leftrightarrow g_0(x) - p^* \geq 0 \, \forall x \in \mathbb{K}$, i.e.,

$g_0(x) - p^*$ **is a nonnegative polynomial on** $\mathbb{K}$.

$\Rightarrow$ Characterize these polynomials ....

$\rightarrow$ (real) algebraic geometry

But we also have

$$p^* \;=\; \min_{\mu} \{ \int g_0(x)\, \mu(dx) \,|\, \mu \in \mathcal{P}(\mathbb{K}) \},$$

where $\mathcal{P}(\mathbb{K})$ is the space of probability measures with support contained in $\mathbb{K}$. Indeed,

$$\int g_0(x)\, \mu(dx) \;\geq p^*, \; \forall \mu \in \mathcal{P}(\mathbb{K}), \tag{0.1}$$

and with $\mu := \delta_{x^*}$ at a global minimizer x^*,

$$\int g_0(x)\, \delta_{x^*}(dx) \;=\; p^*.$$

Observe that both properties are valid for global optima only. Moreover, (0.1) is a linear optimization problem.

$$\min_{\mu} \{ \langle g_0, \mu \rangle \,|\, \langle 1\!\mathrm{l}_{\mathbb{K}}, \mu \rangle \;=\; 1;\; \langle 1\!\mathrm{l}_{\mathbb{K}^c}, \mu \rangle \;=\; 0;\; \mu \;\geq\; 0 \}.$$

$\Rightarrow$ characterize these measures μ ..... (functional analysis)

II. The point of view of moments

The dual linear program of (0.1) is

$$\max_{\gamma,\lambda}\{\gamma \mid \lambda 1\!\mathrm{l}_{\mathbb{K}^c} + \gamma 1\!\mathrm{l}_{\mathbb{K}} \leq\ g_0(x),\ \forall x \in \mathbf{R}^n\},$$

or, equivalently, $\max_\gamma\{\gamma \mid g_0(x) - \gamma\ \geq\ 0\ \text{ on } \mathbb{K}\}$.

Writing

$$\int g_0(x)\,\mu(dx)\ =\ \sum_\alpha (g_0)_\alpha \int x^\alpha\,\mu(dx)\ =\ \sum_\alpha (g_0)_\alpha y_\alpha,$$

with y_α being a moment of order α of μ, (0.1) reads

$$\begin{cases} \min_y \sum_\alpha (g_0)_\alpha y_\alpha \\ \\ y_\alpha\ =\ \int x^\alpha\,\mu(dx)\ \ \forall\alpha\ \text{ for some probability } \mu \text{ on } \mathbb{K} \end{cases}$$

Hence, translate the condition

there is some probability μ on $\mathbb{K}$ such that

$$y_\alpha\ =\ \int x^\alpha\,\mu(dx),\ \forall\alpha \leq r,$$

into a condition on the vector y. This the **$\mathbb{K}$-moment problem**, which dates back to Hausdorff, Markov, Stieltjes, Hamburger, etc ...

- $\mathbb{K} = \mathbf{R}$ **(truncated) Hamburger problem**
- $\mathbb{K} = (\mathbf{R})^+$ **(truncated) Stieltjes problem**
- $\mathbb{K} = [a, b]$ **(truncated) Hausdorff problem**

In the one-dimensional case, there exist **Necessary and sufficient conditions** in terms of **positive semidefinite constraints** on related Hankel matrices $H(y)$...

Particular case of $\mathbb{K} = [0,1]^n$

Hausdorff moment conditions

Ex: $n = 2$; Given a measure $\mu(d(x,z))$ on $\mathbf{R}^2$, let

$$\int x^i(1-x)^m z^j(1-z)^p \, d\mu \;=\; \sum_{k=0}^{m}\sum_{l=0}^{m}\binom{m}{k}\binom{p}{l} x^{i+k}y^{j+l}\, d\mu.$$

Then, given a vector $y \in \mathbf{R}^\infty$, there exists a measure μ on $[0,1]^2$ with

$$\int x^i z^j \, d\mu \;=\; y_{i,j} \qquad \forall i,j = 0,1,\ldots$$

if and only if

$$\sum_{k=0}^{m}\sum_{l=0}^{p}\binom{m}{k}\binom{p}{l} y_{i+k,j+l} \;\geq\; 0,$$

for all $m,p = 1,2,\ldots$ and all $i,j = 0,1,\ldots$
(due to Hausdorff, Bernstein).

Hence, the Hausdorff moment conditions are linear constraints on the y_α's.

... BUT ... notice the **large binomial coefficients** involved ....

II. The point of view of positive polynomials

=Hilbert's 17th problem on the representation of positive polynomials. In the one dimensional case,

$$p(x) \geq 0 \Leftrightarrow p(x) = \sum_{k=1}^{s} q_k(x)^2.$$

Not true anymore in $\mathbf{R}^n$ ....

Representation of polynomials,

positive on $\mathbb{K} := \{x \in \mathbf{R}^n \,|\, g_k(x) \geq 0,\ k = 1, \ldots m\}$

Theorem : [Schmüdgen, Putinar, Jacobi and Prestel]

Assume there is a polynomial $u(x) : \mathbf{R}^n \to \mathbf{R}$ *such that*

$$u(x) = q(x) + \sum_{k=1}^{m} g_k(x) v(x),$$

for some polynomials $q(x), v(x)$ *both sums of squares, and such that* $\{u(x) \geq 0\}$ *is compact. Then:*

Every polynomial, $p(x) : \mathbf{R}^n \to \mathbf{R}$*, strictly positive on* $\mathbb{K}$ *has the representation:*

$$p(x) = \sum_{j=1}^{r_0} q_j(x)^2 + \sum_{k=1}^{m} g_k(x) \sum_{j=1}^{r_k} t_{kj}(x)^2, \tag{0.2}$$

for some (finite) family of polynomials $\{q_j(x)\}$*,* $\{t_{kj}(x)\}$*.*

For instance, the representation (0.2) holds whenever $\{g_k(x) \geq 0\}$ is compact for some k, when all the $g_k(x)$ are linear and $\mathbb{K}$ is compact, etc... In practice, one may also add the redundant constraint $M - \sum_i x_i^2 \geq 0$ for M large enough. This is also the case when one has integrality constraints $x_i^2 = x$ for all i. $\Rightarrow$ **very general result!**

The case of a convex polytope

$$\mathbb{K} := \{x \in \mathbf{R}^n \,|\, Ax \;\leq\; b\}$$

for some matrix $A \in \mathbf{R}^{m \times n}$.

Theorem : [Cassier (1984), Handelman]

Every polynomial, $p(x) : \mathbf{R}^n \to \mathbf{R}$*, strictly positive on* $\mathbb{K}$ *has the representation:*

$$p(x) \;=\; \sum_{|\alpha| \leq s} c_\alpha \, (b - Ax)_1^{\alpha_1} (b - Ax)_2^{\alpha-2} \cdots (b - Ax)_m^{\alpha_m}$$

for some integer s *and nonnegative coefficients* $\{c_\alpha\}$.

Notice the **exponential** number of terms, in contrast to the "linear" Schmüdgen-Putinar representation in terms of squares.

SDP-relaxations

Moment matrix.

Let $(1, \{y\}) \in \mathbf{R}^{s(2r)}$. With $\alpha \in \mathbb{N}^n$, and $|\alpha| = \sum_i \alpha_i$,

$$y_{\alpha_1,\ldots,\alpha_n} \rightsquigarrow \int x_1^{\alpha_1} \cdots x_n^{\alpha_n} \, d\mu.$$

For instance, in $\mathbf{R}^2$ and with $r = 2$, and in the basis of monomials, the **moment matrix** $M_r(y)$ reads

$$M_2(y) = \begin{bmatrix} 1 & | & y_{1,0} & y_{0,1} & | & y_{2,0} & y_{1,1} & y_{0,2} \\ & - & - & - & - & - & - & - \\ y_{1,0} & | & y_{2,0} & y_{1,1} & | & y_{3,0} & y_{2,1} & y_{1,2} \\ y_{0,1} & | & y_{1,1} & y_{0,2} & | & y_{2,1} & y_{1,2} & y_{0,3} \\ & - & - & - & - & - & - & - \\ y_{2,0} & | & y_{3,0} & y_{2,1} & | & y_{4,0} & y_{3,1} & y_{2,2} \\ y_{1,1} & | & y_{2,1} & y_{1,2} & | & y_{3,1} & y_{2,2} & y_{1,3} \\ y_{0,2} & | & y_{1,2} & y_{0,3} & | & y_{2,2} & y_{1,3} & y_{0,4} \end{bmatrix}$$

In general, if $M_r(y)(i,1) = y_\alpha$ and $M_r(y)(1,j) = y_\beta$ then

$$M_r(y)(i,j) = y_{\alpha+\beta} = y_{\alpha_1+\beta_1,\ldots,\alpha_n+\beta_n}$$

Localizing matrix.

Given a polynomial $\theta : \mathbf{R}^n \to \mathbf{R}$ of degree w, with coefficient vector $\theta \in \mathbf{R}^{s(w)}$, let $M_r(\theta y)$ be the **localizing matrix**

$$M_r(\theta y)(i,j) := \sum_\alpha \theta_\alpha y_{\{\alpha(i,j)+\alpha\}}.$$

For instance, with $x \mapsto \theta(x) = 1 - x_1^2 - x_2^2$, $M_2(\theta y) =$

$$\begin{bmatrix} 1 - y_{20} - y_{02}, & y_{10} - y_{30} - y_{12}, & y_{01} - y_{21} - y_{03} \\ y_{10} - y_{30} - y_{12}, & y_{20} - y_{40} - y_{22}, & y_{11} - y_{21} - y_{12} \\ y_{01} - y_{21} - y_{03}, & y_{11} - y_{21} - y_{12}, & y_{02} - y_{22} - y_{04} \end{bmatrix}.$$

If $M_r(y)(i,j) = y_\beta$ then

$$M_r(\theta y)(i,j) = \sum_\alpha \theta_\alpha y_{\beta+\alpha}$$

that is,

$$Mr(\theta y)(i,j) \rightsquigarrow \int x^\beta \, \theta(x) \, \mu(dx)$$

If $(1, y)$ is the vector of moments up to order $2r$ of some probability measure μ on the Borel sets of $\mathbf{R}^n$, then for every polynomial $q(x) : \mathbf{R}^n \to \mathbf{R}$ of degree at most r,

$$\langle q, M_r(y)q \rangle \; = \; \int q(x)^2 \, \mu(dx),$$

so that $M_r(y) \succeq 0$. Similarly,

$$\langle q, M_r(\theta y)q \rangle \; = \; \int \theta(x)q(x)^2 \, \mu(dx),$$

and thus $M_r(\theta y) \succeq 0$ whenever μ is supported on $\{\theta(x) \geq 0\}$.

The theory of moments identifies those vectors y with $M_r(y) \succeq 0$ that are the moments of some measure μ.

The $\mathbb{K}$-moment problem identifies those vectors y with $M_r(y) \succeq 0$ that are moment of a measure μ with support contained in $\mathbb{K}$.

Dual theory in algebraic geometry of representation of polynomials, positive on a semi-algebraic set $\mathbb{K}$

Introduce the family $\{\mathbb{Q}_i\}$ of SDP-relaxations

$$\mathbb{Q}_i \begin{cases} \min_y \sum_\alpha (g_0)_\alpha y_\alpha & \\ M_i(y) & \succeq 0 \\ M_{i-v_k}(g_k y) & \succeq 0, \quad k = 1, \ldots m. \end{cases}$$

and the family $\{(\mathbb{Q}_i)^*\}$ of their dual

$$\mathbb{Q}_i^* \begin{cases} \max_{X, Z_1, \ldots Z_m \succeq 0} -X(1,1) - \sum_{k=1}^m g_k(0) Z_k(1,1) \\ \text{s.t. } \langle X, B_\alpha \rangle + \sum_{k=1}^m \langle Z_k, C_\alpha^k \rangle = (g_0)_\alpha, \ \forall \alpha. \end{cases}$$

where we write

$$\begin{aligned} M_i(y) &= \sum_\alpha y_\alpha B_\alpha \\ M_{i-v_k}(g_k y) &= \sum_\alpha y_\alpha C_\alpha^k, \ k = 1, \ldots m, \end{aligned}$$

Interpretation :

From the dual, $\max \mathbb{Q}_i^* \leq p^*$ and

$$p(x) - \max \mathbb{Q}_i^* = \sum_{j=1}^s q_j(x)^2 + \sum_{k=1}^m g_k(x) \left[\sum_{l=1}^{s_k} q_{kl}(x)^2 \right],$$

with degree$(q_i) \leq i$ and degree$(q_{kl}) \leq i - v_k$.

SDP-relaxation $\longleftrightarrow$ Schmüdgen-Putinar representation of $p(x) - p^*$.

Theorem 1. *Assume that there exists a polynomial* $u(x) : \mathbf{R}^n \to \mathbf{R}$ *of the form (0.2) with* $\{u(x) \geq 0\}$ *compact. Then:*

$$\inf \mathbb{Q}_i \uparrow p^* = \min \mathbb{P}. \tag{0.3}$$

In addition, if $p(x) - p^*$ *has the representation (0.2) for polynomials* $\{q_j(x)\}$ *et* $\{t_{kj}(x)\}$*, of degree at most* N*, then*

$$p^* = \min \mathbb{Q}_i, \qquad \forall i \geq N, \tag{0.4}$$

and for every optimal solution x^* *of* $\mathbb{P}$*, the vector*

$$y^* := (x_1^*, \ldots, x_n^*, \ldots (x_1^*)^{2i}, \ldots (x_n^*)^{2i}) \tag{0.5}$$

is an optimal solution of $\mathbb{Q}_i$*.*

Karush-Kuhn-Tucker Global optimality conditions

Proposition 2. *Assume that there exists a polynomial* $u(x):\mathbf{R}^n\to\mathbf{R}$ *of the form (0.2) with* $\{u(x)\geq 0\}$ *compact, and let* x^* *be an optimal solution of* $\mathbb{P}$. *If* $p(x)-p^*$ *has the representation (0.2), then*

$$g_k(x^*)\left[\sum_{j=1}^{r_k} t_{kj}(x^*)^2\right] \;=\; 0,\; k=1,\dots m \tag{0.6}$$

$$\nabla g_0(x^*) \;=\; \sum_{k=1}^{m}\nabla g_k(x^*)\left[\sum_{j=1}^{r_k} t_{kj}(x^*)^2\right] \tag{0.7}$$

Thus, we may interpret the representation (0.2) of $p(x)-p^*$ as a **global optimality condition** à la Karush-Kuhn-Tucker, with **polynomials multipliers** $\sum_{j=1}^{r_k} t_{kj}(x)^2$ in lieu of the usual scalars λ_k^*, $k=1,\dots m$.

Moreover, if (x^*,λ^*) is KKT optimal point of $\mathbb{P}$, and the gradients $\{\nabla g_k(x^*)\}$ are linearly independent, then

$$\sum_{j=1}^{r_k} t_{kj}(x)^2 \;=\; \lambda_k^*,\ \forall k=1,\dots m.$$

LP-Relaxations

Basic idea: [Shor, Sherali and Adams]

(1) add **redundant** constraints of the form

$$g_1(x)^{\alpha_1} g_2(x)^{\alpha_2} \cdots g_m(x)^{\alpha_m} \;\geq\; 0,$$

with $|\alpha| \;\leq\; \delta$, fixed.

(2) **linearize** all the terms

$$y \;\mapsto\; y_\beta \;:=\; x_1^{\beta_1} x_2^{\beta_2} \cdots x_n^{\beta_n},$$

to replace the nonlinear constraint

$$g_1(x)^{\alpha_1} \cdots g_m(x)^{\alpha_m} \;\geq\; 0,$$

by the linear constraint

$$\sum_\beta c_\beta y_\beta \;\geq\; b$$

The LP-Relaxation of order δ is the LP pogram

$$\mathbb{P}_\delta \to \min_{y_\beta} \{\, c'_\beta y_\beta \mid A_\delta y \;\geq\; b_\delta \},$$

with dual

$$\mathbb{P}^*_\delta \to \max_{\lambda \geq 0} \{\, b'_\delta \lambda \mid A'_\delta \lambda \;=\; c_\delta \}$$

Interpretation

Let $\rho_\delta := \min \mathbb{P}_\delta = \max \mathbb{P}^*_\delta$. Then : $\rho_\delta \leq p^*$ and :

$$g_0(x) - \rho_\delta = \sum_{|\alpha| \leq \delta} \lambda_\alpha g_1(x)^{\alpha_1} g_2(x)^{\alpha_2} \cdots g_m(x)^{\alpha_m}, \; x \in \mathbf{R}^n$$

The dual aims at representing $g_0(x) - p^*$ as in **Cassier, Handelman**'s representation, which in principle is valid only when the g_k are linear and define a polytope $\mathbb{K}$.

Corollary : *Let $\mathbb{K}$ be a convex polytope, i.e., the g_k's are linear : Then*

$$\rho_\delta \uparrow p^* \qquad \textit{as } |\delta| \to \infty,$$

that is, the LP-relaxations converge.

Remarks: (i) ill-conditioned when δ large.

Example : $\mathbb{K} = [0,1]^n$. From the constraints $x \geq 0$ and $(1 - x_i) \geq 0$, the linear constraints coming from the linearization of $\prod_{i \in I} x_i^{\alpha_i} \prod_{j \in J} (1 - x_j)^{\alpha_j}$ contains **large binomial coefficients**.

They are in fact the **Hausdorff moment conditions** on y to be the vector of moments of a probability measure on $[0,1]^n$.

Let $I(x)$ be the set of active constraints at a feasible point x, i.e., $i \in I(x) \Rightarrow g_i(x) = 0$.

(ii) No relaxation can be exact if a global minimum is in the interior of $\mathbb{K}$ or if there is a nonoptimal feasible point x with $I(x) = I(x^*)$.

Conclusion

SDP-relaxations $\mathbb{Q}_i$	**LP-relaxations** $\mathbb{P}_i$
LMI constraints	**Linear constraints**
$p(x) - \min \mathbb{Q}_i = \Sigma^2 + \sum_{k=1}^m g_k(x)\Sigma^2$	$p(x) - \min \mathbb{P}_i = \sum_\alpha g_1^{\alpha_1} \cdots g_m^{\alpha_m}$
Schmüdgen-Putinar	**Cassier, Handelman** (polytope)
$0-1 \to \min \mathbb{Q}_{n+v} = p^*$	$0-1 \to \min \mathbb{P}_n = p^*$
$\mathbb{K}$ **compact**. $\min \mathbb{Q}_i \uparrow p^*$	$\mathbb{K}$ **polytope**. $\min \mathbb{P}_i \uparrow p^*$
$n=1$; a single relaxation	$n=1$; $\min \mathbb{P}_i \uparrow p^*$
SDP-packages : limited	LP-packages : unlimited
	ill-conditionning; binomial coeff.

Bounds on measures with moment conditions [Lasserre (2002)], Annals Appl. Prob.

Let $\Gamma \subset \mathbb{N}^n$ and let $\{\gamma_\alpha\}_{\alpha\in\Gamma}$ be a given finite sequence of scalars. **Problem:** Given a semi-algebraic set

$$\mathbb{K} := \{x \in \mathbf{R}^n \,|\, g_i(x) \geq 0, \quad i = 1, \ldots, m\}$$

we want to find, or approximate

$$\rho^* \; := \; \sup_{\mu \in \mathcal{P}(\mathbf{R}^n)} \{\mu(\mathbb{K}) \,|\, \int x^\alpha \, d\mu \; = \; \gamma_\alpha \quad \alpha \in \Gamma\}$$

Applications in Probability, Finance, Queuing, ...

Work motivated by nice results from **Bertsimas and Popescu** (1999)-(2002) for the case $\mathbb{K}$ convex, and first- and second order moment conditions (see also further extensions).

Write $\mu = \varphi + \psi$ with $\varphi(K^c) = 0$ and let $\{y_\alpha, z_\alpha\}$ be the respective moments of φ, ψ, so that

$$\int x^\alpha d\mu \; = \; y_\alpha + z_\alpha \quad \alpha \in \mathbb{N}^n.$$

$$\rho^* \; = \; \left\{ \begin{array}{rl} & \sup y_0 \\ \text{s.t.} & y_\alpha + z_\alpha \; = \gamma_\alpha \quad \alpha \in \Gamma \end{array} \right.$$

SDP-relaxation

$$\mathbb{Q}_r \begin{cases} \quad\;\; \sup y_0 \\ \text{s.t.}\;\; M_r(y), M_r(z) \succeq 0 \\ \quad\;\; M_{r-v_i}(g_i y) \succeq 0 \quad i = 1, \ldots, m \\ \quad\;\; y_\alpha + z_\alpha = \gamma_\alpha \quad \alpha \in \Gamma \end{cases}$$

Theorem: (a) As $r \to \infty$

$$\sup \mathbb{Q}_r \downarrow \delta \geq \rho^*.$$

(b) If in addition, $\mathbb{Q}_r$ is solvable and an optimal solution y^*, z^* satisfies

$$\text{rank} M_r(y^*) = \text{rank} M_{r-d}(y^*); \quad \text{rank} M_r(z^*) = \text{rank} M_{r-1}(z^*),$$

then

$$\max \mathbb{Q}_r = \rho^* = \max_{\mu \in \mathcal{P}(\mathbf{R}^n)} \{\mu(\mathbb{K}) \mid \int x^\alpha \, d\mu = \gamma_\alpha, \quad \alpha \in \Gamma.$$

Invariant prob. measures of Markov chains

Let $(X, \mathcal{B}, Q)$ be a time-homogeneous Markov chain $\Phi_\bullet = (\Phi_0, \Phi_1, \ldots)$ with **state space** X, and **t.p.f.** Q.

- $Q(x, .)$ is a **prob. measure** on $\mathcal{B}$ for all $x \in X$.
- $x \mapsto Q(x, B)$ is **measurable** for all $B \in \mathcal{B}$.

$$\mathrm{Prob}(\Phi_t \in B \,|\, \Phi_{t-1} = x) \;=\; Q(x, B) \quad B \in \mathcal{B},\; x \in X.$$

$\mu \in \mathcal{P}(X)$ is an invariant prob. measure for the MC $\Phi_\bullet$ if

$$\mu(B) \;=\; \int_X P(x, B)\, \mu(dx) \qquad \forall\, B \in \mathcal{B}.$$

If μ is unique and $f \in L_1(\mu)$ then :

$$\lim_{T \to \infty} E_x \frac{1}{T} \sum_{t=0}^{T-1} f(\Phi_t) \;=\; \int_X f\, d\mu \quad \mu\text{−a.e.},$$

and for μ-a.a. $x \in X$,

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} f(\Phi_t) \;=\; \int_X f\, d\mu \quad P_x\text{−a.s}$$

Simulation gives only an **estimate** of $\int_X f\, d\mu$.

Remark: $X \equiv \mathbf{R}^n$. If $Qf \in \mathbf{R}[x]$ when $f \in \mathbf{R}[x]$ then

$$y_\alpha = \int x^\alpha \, d\mu \;=\; \int (Qx^\alpha)\, d\mu \;=\; \langle A_\alpha, y \rangle$$

SDP-relaxations: Let $f = \sum_\alpha f_\alpha x^\alpha \in \mathbf{R}[x]$.

$$\mathbb{Q}_r \left\{ \begin{array}{lll} & \max_y \text{ (or } \min_y) & \sum_\alpha f_\alpha y_\alpha \\ \text{s.t.} & M_r(y) & \succeq 0 \\ & y_\alpha - \langle A_\alpha, y\rangle & = 0, \quad \forall\, |\alpha| \leq r \end{array} \right.$$

$$\min \mathbb{Q}_r \;\leq\; \int_X f\, d\mu \;\leq\; \max \mathbb{Q}_r \qquad \forall\, r$$

so that

$$\sup_r \min \mathbb{Q}_r \;=\; \underline{\rho} \leq \int_X f\, d\mu \;\leq\; \overline{\rho} \;=\; \inf_r \max \mathbb{Q}_r$$

If μ is the unique inv. prob. measure then one gets upper and lower bounds on $\int f\, d\mu$. If not, then :

$$\sup_r \min \mathbb{Q}_r \;=\; \underline{\rho} \;\leq\; \inf_{\mu Q = \mu} \int_X f\, d\mu$$

and

$$\sup_{\mu Q = \mu} \int_X f\, d\mu \;\leq\; \overline{\rho} \;=\; \inf_r \max \mathbb{Q}_r$$

LP-relaxations: Stockbridge, Helmes

$$\mathrm{LP}_r \left\{ \begin{array}{ll} & \max_y \text{ (or } \min_y) \sum_\alpha f_\alpha y_\alpha \\ \text{s.t.} & \text{Hausdorff moment conditions} \\ & y_\alpha - \langle A_\alpha, y\rangle \;=\; 0, \quad \forall\, |\alpha| \leq r \end{array} \right.$$

SDP-relaxations better than LP-relaxations

I. Barnsley Iterated function systems:

$$x_{t+1} = f_{\xi_t}(x_{t-1}), \quad t = 1, \ldots,$$

- $X \equiv \mathbf{R}^n$, $\xi_\bullet$ is a Markov chain on a finite set $S = \{1, \ldots, m\}$ with **t.p.f.** $P(\xi_t = j | \xi_{t-1} = i) = p_{ij}$ and unique invariant prob. distribution $\gamma = (\gamma_1, \ldots, \gamma_m)$.
- $f_\xi : \mathbf{R}^n \to \mathbf{R}$ is a polynomial

Evaluate or approximate

$$\rho^* := \lim_{T\to\infty} E_\nu \left[\frac{1}{T} \sum_{t=0}^{T-1} h(x_t) \right],$$

when $h := 1\!\!1_{\mathbb{K}}$ or $h \in \mathbf{R}[x_1, \ldots, x_n]$, and $\nu \in \mathcal{P}(\mathbf{R}^n)$ is the initial prob. distribution of $x_0 \in \mathbf{R}^n$.

$\{X_t\}$ is a time-homogeneous, $\mathbf{R}^n$-valued Markov chain. Usually ρ^* is **estimated** via **Simulation**. The result depends on the inital distribution if there are several invariant prob. distributions.

Let μ be an invariant prob. distribution and assume that all its moments $\{y_\alpha\}$ are finite.

$$y_\alpha \,=\, \int_{\mathbf{R}^n} x^\alpha d\mu \,=\, \int_{\mathbf{R}^n} \sum_{j \in S} \gamma_j f_j(x)^\alpha \, d\mu \,=\, \langle A_\alpha, y \rangle$$

Example: Logistic map:

$X \equiv [0,1]$ and $x_{t+1} = 4x_t(1-x_t)$.

$$y_n \,=\, \int_0^1 x^n \, d\mu \,=\, [4x(1-x)]^n \, d\mu \,=\, \sum_{j=0}^{2n} \beta_{jn} y_j$$

Invariant prob. measures:

- Uncountably many inv. prob. measures with support, all cycles of order 1, 2, etc ...

- 1 inv. prob. measure, absolutely continuous w.r.t. λ, with density $x \mapsto [\,\pi\sqrt{x(1-x)}\,]^{-1}$

II. Diffusions

Exit times. Consider a continuous Markov process on $\Phi = \{\Phi_t, t \geq 0\}$ on $\mathbf{R}^n$, with infinitesimal generator $\mathcal{A}$ (with domain $\mathcal{D}$).

Let $E_0 \subset \mathbf{R}^n$ be a bounded open set and E_0^c its complement. Let τ be the **first time** the process Φ **hits** E_0^c.

Problem: Evaluate or approximate the distribution of τ, e.g., its moments $E[\tau]$, $E[\tau^n]$, ...

Let μ_0, μ_1 be the occupation measure and the exit time distribution,

$$\mu_0(B) = E \int_0^\tau 1\!\!1_B(\Phi_s)\, ds; \quad \mu_1(B) = \text{Prob}(X_\tau \in B) \quad B \in \mathcal{B}.$$

It follows that

$$** \quad f(x_0) + \int_{E_0} \mathcal{A} f(x) \mu_0(dx) - \int_{E_0^c} f(x)\, \mu_1(dx) = 0.$$

In general, we have $\mathcal{A} = \partial/\partial x_i, \partial^2/\partial x_i \partial x_j, \ldots$ so that $\mathcal{A} f \in \mathbf{R}[x]$ if $f \in \mathbf{R}[x]$.

Then (**) generates **linear constraints** on the **moments** of μ_0, μ_1!

Example: Cox-Ingersoll-Ross interest rate model (Stockbridge and Helmes).

$$dY_t = (aY_t + \beta)dt + \sigma\sqrt{Y_t}\,dW_t,$$

with $Y(0) \in (r,1)$, $\alpha \in \mathbf{R}, \beta, \sigma \geq 0$ and $0 < r < 1$.

$$\mathcal{A}f(x) = \frac{\sigma^2}{2} x f"(x) + (\alpha x + \beta) f'(x),$$

with f twice continuously differentiable.

Let τ be the first exit time of $\{Y_t\}$ from $(r,1)$, so that $E_0 = (r,1)$ and $E_0^c = \{r\} \cup \{1\}$. So μ_0 has support on $(r,1)$ and $\mu_1 = (\mu_1(\{r\}), \mu_1(\{1\})) = (p,q)$ with $p+q = 1$.

Let $\{y_\alpha\}$ and $\{z_\alpha\}$ be the moments of μ_0 and μ_1.

$$z^\alpha = pr^\alpha + q \qquad \alpha \in \mathbb{N}.$$

When $f = x^\alpha$, the equation

$$f(x_0) + \int_{E_0} \mathcal{A}f(x)\mu_0(dx) - \int_{E_0^c} f(x)\,\mu_1(dx) = 0.$$

define **linear constraints**

$$\langle U_\alpha, y \rangle + pr^\alpha + q = 0, \quad \alpha \in \mathbb{N},$$

on the moments y and the scalars p, q

As $E[\tau] = \mu_0(X)$, the SDP-relaxation reads:

$$\mathbb{Q}_m \left\{ \begin{array}{lcl} \max\ \text{(or, min)}\ y_0 & & \\ M_r(y) & \succeq 0 & \\ \langle U_\alpha, y\rangle - pr^\alpha - q & = \ 0 & \quad \forall\,|\alpha| \leq m \\ p, q \geq 0,\ p + q & = \ 1 & \end{array} \right.$$

Again

$$\sup_m \min \mathbb{Q}_m \ \leq \ E[\tau] \ \leq \ \inf_m \max \mathbb{Q}_m,$$

and very good results are reported by Stockbridge and Helmes for the LP-relaxations (which are less efficient than SDP-relaxations)

Extension to higher moments $E[\tau^n]$ easy by using an extended generator for the joint process $\{(X_t, t), t \geq 0\}$.