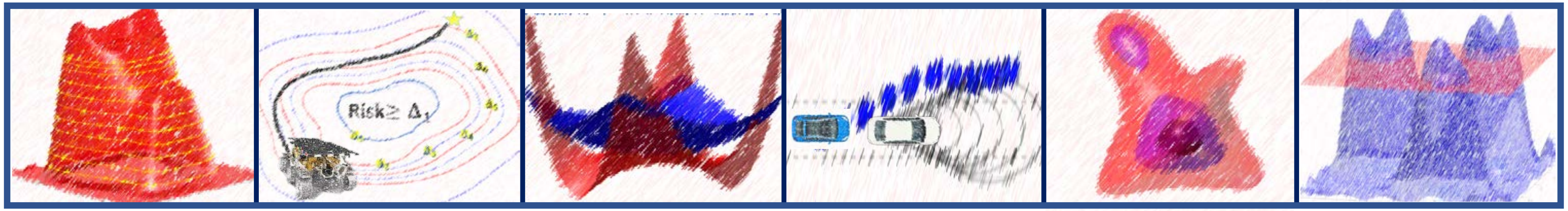
Lecture 4

Measure and Moments For Nonlinear Optimization

MIT 16.S498: Risk Aware and Robust Nonlinear Planning

Fall 2019

Ashkan Jasour



Nonlinear (nonconvex) Optimization

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & f(x) \\ \text{subject to} & g_i(x) \geq 0, \;\; i = 1, \ldots, m \end{array}$$

Objective function and constraints are polynomial functions.

Goal: Find Convex Relaxations of Nonlinear Optimization

Tools:

i) Measure (e.g. probability distribution) and Moments

ii) Semidefinite Programs (SDP)

Nonlinear (nonconvex) Optimization

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \geq 0, \;\; i = 1, \ldots, m \end{aligned}$$

Tools: i) Measure and Moments ii) Semidefinite Programs

Step 1: Infinite-dimensional LP

Reformulate **Nonlinear Optimization** problem in terms of **Probability Distributions** (**measures**)

Step 2: Infinite-dimensional SDP

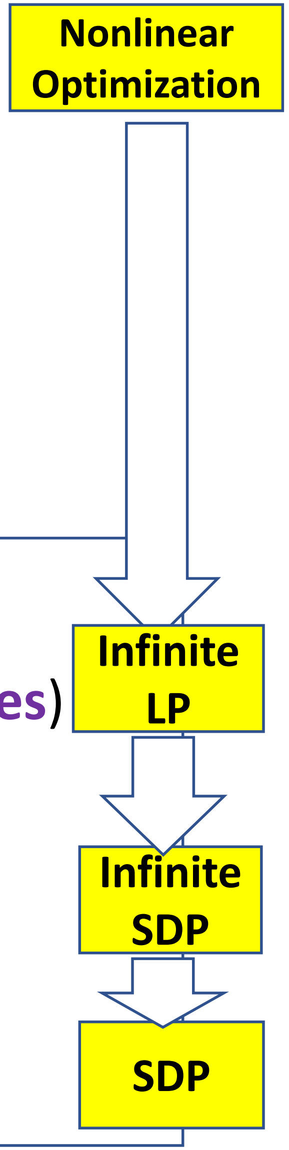
Reformulate **Problem of Step 1** in terms of **Moments (higher order statistics)**

Step 3: Finite SDP

Truncate matrices of **SDP of Step 2 (truncated moments)**

Nonlinear (nonconvex) Optimization

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \geq 0, \;\; i = 1, \ldots, m \end{aligned}$$



Tools: i) Measure and Moments ii) Semidefinite Programs

Step 1: Infinite-dimensional LP

Reformulate **Nonlinear Optimization** problem in terms of **Probability Distributions** (**measures**)

Step 2: Infinite-dimensional SDP

Reformulate **Problem of Step 1** in terms of **Moments (higher order statistics)**

Step 3: Finite SDP

Truncate matrices of **SDP of Step 2 (truncated moments)**

Overview

1 Nonlinear Optimization:

- **Decision variables:** $x \in \mathbb{R}^n$
- Look for $x \in \Omega$ to minimize the $p(x)$
  - Unconstrained Optimization $\Omega = \mathbb{R}^n$
  - Constrained Optimization $\Omega = \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ i = 1, ..., m\}$

$$\mathrm{P}^* = \underset{x \in \Omega}{\text{minimize}} \quad p(x)$$

1 Nonlinear Optimization:

- **Decision variables:** $x \in \mathbb{R}^n$
- Look for $x \in \Omega$ to minimize the $p(x)$
  - Unconstrained Optimization $\Omega = \mathbb{R}^n$
  - Constrained Optimization $\Omega = \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ i = 1, ..., m\}$

$$\mathrm{P}^* = \underset{x \in \Omega}{\text{minimize}} \quad p(x)$$

Step 1

2 Optimization in terms of Probability distributions (measures):

- **Decision variable:** $pr(x)$ probability distribution of x (treat x as a random variable)
- Look for $pr(x)$ to minimize the $\mathrm{E}[p(x)]$

1 Nonlinear Optimization:

- **Decision variables:** $x \in \mathbb{R}^n$
- Look for $x \in \Omega$ to minimize the $p(x)$
  - Unconstrained Optimization $\Omega = \mathbb{R}^n$
  - Constrained Optimization $\Omega = \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0, \ i = 1, \ldots, m\}$

$$\mathrm{P}^* = \underset{x \in \Omega}{\text{minimize}} \quad p(x)$$

Step 1

2 Optimization in terms of Probability distributions (measures):

- **Decision variable:** $pr(x)$ probability distribution of x (treat x as a random variable)
- Look for $pr(x)$ to minimize the $\mathrm{E}[p(x)]$

$$\underset{\substack{pr(x) \in \text{space of probability distributions} \\ \text{defined on } \Omega}}{\text{minimize}} \quad \mathrm{E}[p(x)] = \int p(x) pr(x) dx$$

① Nonlinear Optimization:

- **Decision variables:** $x \in \mathbb{R}^n$
- Look for $x \in \Omega$ to minimize the $p(x)$
  - Unconstrained Optimization $\Omega = \mathbb{R}^n$
  - Constrained Optimization $\Omega = \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0, \ i = 1, ..., m\}$

Optimal Solution: x^*

Optimal Objective: $p(x^*)$

$$\mathrm{P}^* = \underset{x \in \Omega}{\text{minimize}} \quad p(x)$$

Step 1

② Optimization in terms of Probability distributions (measures):

- **Decision variable:** $pr(x)$ probability distribution of x (treat x as a random variable)
- Look for $pr(x)$ to minimize the $\mathrm{E}[p(x)]$

$$\underset{\substack{pr(x) \in \text{space of probability distributions} \\ \text{defined on } \Omega}}{\text{minimize}} \quad \mathrm{E}[p(x)] = \int p(x) pr(x) dx$$

1 Nonlinear Optimization:

- **Decision variables:** $x \in \mathbb{R}^n$
- Look for $x \in \Omega$ to minimize the $p(x)$
  - Unconstrained Optimization $\Omega = \mathbb{R}^n$
  - Constrained Optimization $\Omega = \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ i = 1, \ldots, m\}$

Optimal Solution: x^*

Optimal Objective: $p(x^*)$

$$\mathrm{P}^* = \underset{x \in \Omega}{\text{minimize}} \quad p(x)$$

Step 1

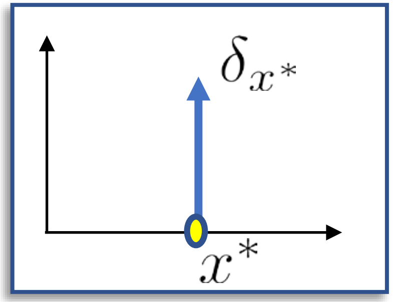
2 Optimization in terms of Probability distributions (measures)

- **Decision variable:** $pr(x)$ probability distribution of x (treat x as
- Look for $pr(x)$ to minimize the $\mathrm{E}[p(x)]$

$$\underset{\substack{pr(x) \in \text{space of probability distributions} \\ \text{defined on } \Omega}}{\text{minimize}} \quad \mathrm{E}[p(x)] = \int p(x) pr(x) dx$$

Optimal Solution:

δ_{x^*}: Dirac probability at x^*



Optimal Objective:

$$\mathrm{E}[p(x)] = \int p(x) \delta_{x^*} dx = p(x^*)$$

1 Nonlinear Optimization:

- **Decision variables:** $x \in \mathbb{R}^n$
- Look for $x \in \Omega$ to minimize the $p(x)$
  - Unconstrained Optimization $\Omega = \mathbb{R}^n$
  - Constrained Optimization $\Omega = \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ i = 1, \ldots, m\}$

$$\mathrm{P}^* = \underset{x \in \Omega}{\text{minimize}} \quad p(x)$$

- Nonlinear in x

Given x_1, x_2 :

$$p(x_1 + x_2) \neq p(x_1) + p(x_2)$$

Step 1

2 Optimization in terms of Probability distributions (measures):

- **Decision variable:** $pr(x)$ probability distribution of x (treat x as a random variable)
- Look for $pr(x)$ to minimize the $\mathrm{E}[p(x)]$

$$\underset{\substack{pr(x) \in \text{space of probability distributions} \\ \text{defined on } \Omega}}{\text{minimize}} \quad \mathrm{E}[p(x)] = \int p(x) pr(x) dx$$

1 Nonlinear Optimization:

- **Decision variables:** $x \in \mathbb{R}^n$
- Look for $x \in \Omega$ to minimize the $p(x)$
  - Unconstrained Optimization $\Omega = \mathbb{R}^n$
  - Constrained Optimization $\Omega = \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ i = 1, ..., m\}$

$$\mathrm{P}^* = \underset{x \in \Omega}{\text{minimize}} \quad p(x)$$

- Nonlinear in x

Given x_1, x_2 :

$$p(x_1 + x_2) \neq p(x_1) + p(x_2)$$

Step 1

2 Optimization in terms of Probability distributions (measures)

- **Decision variable:** $pr(x)$ probability distribution of x (treat x as
- Look for $pr(x)$ to minimize the $\mathrm{E}[p(x)]$

$$\underset{\substack{pr(x) \in \text{space of probability distributions} \\ \text{defined on } \Omega}}{\text{minimize}} \quad \mathrm{E}[p(x)] = \int p(x) pr(x) dx$$

- Linear in $pr(x)$

Given $pr_1(x)$, $pr_2(x)$:

$$\int p(x) \left(pr_1(x) + pr_2(x)\right) dx =$$

$$\int p(x) pr_1(x) dx + \int p(x) pr_2(x) dx$$

1 Nonlinear Optimization:

- **Decision variables:** $x \in \mathbb{R}^n$
- Look for $x \in \Omega$ to minimize the $p(x)$
  - Unconstrained Optimization $\Omega = \mathbb{R}^n$
  - Constrained Optimization $\Omega = \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ i = 1, ..., m\}$

Nonlinear Optimization

$$\mathrm{P}^* = \underset{x \in \Omega}{\text{minimize}} \quad p(x)$$

Step 1

2 Optimization in terms of Probability distributions (measures):

- **Decision variable:** $pr(x)$ probability distribution of x (treat x as a random variable)
- Look for $pr(x)$ to minimize the $\mathrm{E}[p(x)]$

Linear Optimization

$$\underset{\substack{pr(x) \in \text{space of probability distributions} \\ \text{defined on } \Omega}}{\text{minimize}} \quad \mathrm{E}[p(x)] = \int p(x) pr(x) dx$$

1 Nonlinear Optimization:

- **Decision variables:** $x \in \mathbb{R}^n$
- Look for $x \in \Omega$ to minimize the $p(x)$
  - Unconstrained Optimization $\Omega = \mathbb{R}^n$
  - Constrained Optimization $\Omega = \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ i = 1, \ldots, m\}$

Nonlinear Optimization

$$\mathrm{P}^* = \underset{x \in \Omega}{\text{minimize}} \quad p(x)$$

Step 1

2 Optimization in terms of Probability distributions (measures):

- **Decision variable:** $pr(x)$ probability distribution of x (treat x as a random variable)
- Look for $pr(x)$ to minimize the $\mathrm{E}[p(x)]$

Infinite dimensional LP

$$\underset{\substack{pr(x) \in \text{space of probability distributions} \\ \text{defined on } \Omega}}{\text{minimize}} \quad \mathrm{E}[p(x)] = \int p(x) pr(x) dx$$

2 Optimization in terms of Probability distributions (measures):

- **Decision variable:** $pr(x)$ probability distribution of x
- Look for $pr(x)$ to minimize the $\mathrm{E}[p(x)]$

$$\underset{\substack{pr(x) \in \text{space of probability distributions} \\ \text{defined on } \Omega}}{\text{minimize}} \quad \mathrm{E}[p(x)] = \int p(x) pr(x) dx$$

Step 2

3 Optimization in terms of Moments

- **Decision variable:** $\text{moments: } y_\alpha = \mathrm{E}[x^\alpha] = \int x^\alpha pr(x) dx$
- Look for sequence of moments $[y_\alpha, \alpha = 0, 1, ..., \infty]$ to minimize the $\mathrm{E}[p(x)]$

② Optimization in terms of Probability distributions (measures):

- **Decision variable:** $pr(x)$ probability distribution of x
- Look for $pr(x)$ to minimize the $\mathrm{E}[p(x)]$

$$\underset{\substack{pr(x)\in \text{space of probability distributions} \\ \text{defined on } \Omega}}{\text{minimize}} \quad \mathrm{E}[p(x)] = \int p(x)pr(x)dx$$

Step 2

③ Optimization in terms of Moments

- **Decision variable:** $\text{moments: } y_\alpha = \mathrm{E}[x^\alpha] = \int x^\alpha pr(x)dx$
- Look for sequence of moments $[y_\alpha, \alpha = 0, 1, ..., \infty]$ to minimize the $\mathrm{E}[p(x)]$

$$\underset{y\in \text{space of moments}}{\text{minimize}} \quad \mathrm{E}[p(x)] = \int p(x)pr(x)dx$$

2 Optimization in terms of Probability distributions (measures):

- **Decision variable:** $pr(x)$ probability distribution of x
- Look for $pr(x)$ to minimize the $\mathrm{E}[p(x)]$

$$\underset{\substack{pr(x)\in \text{space of probability distributions} \\ \text{defined on } \Omega}}{\text{minimize}} \quad \mathrm{E}[p(x)] = \int p(x)pr(x)dx$$

Step 2

3 Optimization in terms of Moments

- **Decision variable:** moments: $y_\alpha = \mathrm{E}[x^\alpha] = \int x^\alpha pr(x)dx$
- Look for sequence of moments $[y_\alpha, \alpha = 0, 1, ..., \infty]$ to minimize the $\mathrm{E}[p(x)]$

$$\underset{y\in \text{space of moments}}{\text{minimize}} \quad \mathrm{E}[p(x)] = \int p(x)pr(x)dx = \int \sum_\alpha p_\alpha x^\alpha pr(x)dx = \sum_\alpha p_\alpha \underbrace{\int x^\alpha pr(x)dx}_{\text{Moment of order } \alpha} = \sum_\alpha p_\alpha y_\alpha$$

polynomial

2 Optimization in terms of Probability distributions (measures):

- **Decision variable:** $pr(x)$ probability distribution of x
- Look for $pr(x)$ to minimize the $\mathrm{E}[p(x)]$

$$\underset{\substack{pr(x)\in \text{space of probability distributions} \\ \text{defined on } \Omega}}{\text{minimize}} \quad \mathrm{E}[p(x)] = \int p(x)pr(x)dx$$

Step 2

3 Optimization in terms of Moments

- **Decision variable:** moments: $y_\alpha = \mathrm{E}[x^\alpha] = \int x^\alpha pr(x)dx$
- Look for sequence of moments $[y_\alpha, \alpha = 0, 1, ..., \infty]$ to minimize the $\mathrm{E}[p(x)]$

$$\underset{\text{vector } y\in \text{space of moments}}{\text{minimize}} \quad \sum_\alpha p_\alpha y_\alpha \longrightarrow \text{Linear objective in terms of moments}$$

2 Optimization in terms of Probability distributions (measures):

- **Decision variable:** $pr(x)$ probability distribution of x
- Look for $pr(x)$ to minimize the $\mathrm{E}[p(x)]$

$$\underset{\substack{pr(x)\in \text{space of probability distributions} \\ \text{defined on } \Omega}}{\text{minimize}} \quad \mathrm{E}[p(x)] = \int p(x)pr(x)dx$$

Step 2

3 Optimization in terms of Moments

- **Decision variable:** $\text{moments: } y_\alpha = \mathrm{E}[x^\alpha] = \int x^\alpha pr(x)dx$
- Look for sequence of moments $[y_\alpha, \alpha = 0, 1, ..., \infty]$ to minimize the $\mathrm{E}[p(x)]$
- Relation between moments and probability distributions : LMI in terms of y

$$\underset{\text{vector } y\in \text{space of moments}}{\text{minimize}} \quad \sum_\alpha p_\alpha y_\alpha \longrightarrow \text{Linear objective in terms of moments}$$

2 Optimization in terms of Probability distributions (measures):

- **Decision variable:** $pr(x)$ probability distribution of x
- Look for $pr(x)$ to minimize the $\mathrm{E}[p(x)]$

$$\underset{\substack{pr(x)\in\text{space of probability distributions}\\ \text{defined on }\Omega}}{\text{minimize}} \quad \mathrm{E}[p(x)] = \int p(x)pr(x)dx$$

Step 2

3 Optimization in terms of Moments

- **Decision variable:** moments: $y_\alpha = \mathrm{E}[x^\alpha] = \int x^\alpha pr(x)dx$
- Look for sequence of moments $[y_\alpha, \alpha = 0, 1, \ldots, \infty]$ to minimize the $\mathrm{E}[p(x)]$
- Relation between moments and probability distributions : LMI in terms of y

$$\underset{\text{vector } y}{\text{minimize}} \sum_\alpha p_\alpha y_\alpha \longrightarrow \text{Linear objective in terms of moments}$$

$$\text{subject to LMI}(y) \longrightarrow y \text{ is moment vector of probability distribution}$$

2 Optimization in terms of Probability distributions (measures):

- **Decision variable:** $pr(x)$ probability distribution of x
- Look for $pr(x)$ to minimize the $\mathrm{E}[p(x)]$

Infinite dimensional LP

$$\underset{\substack{pr(x)\in \text{space of probability distributions} \\ \text{defined on } \Omega}}{\text{minimize}} \quad \mathrm{E}[p(x)] = \int p(x) pr(x) dx$$

Step 2

3 Optimization in terms of Moments

- **Decision variable:** moments: $y_\alpha = \mathrm{E}[x^\alpha] = \int x^\alpha pr(x) dx$
- Look for sequence of moments $[y_\alpha, \alpha = 0, 1, \dots, \infty]$ to minimize the $\mathrm{E}[p(x)]$
- Relation between moments and probability distributions : LMI in terms of y

Infinite dimensional SDP

$$\underset{\text{vector } y}{\text{minimize}} \sum_\alpha p_\alpha y_\alpha \longrightarrow$$ Linear objective in terms of moments

$$\text{subject to } \mathrm{LMI}(y) \longrightarrow$$ y is moment vector of probability distribution

(3) Optimization in terms of Moments

- **Decision variable:** $\text{moments: } y_\alpha = \mathrm{E}[x^\alpha] = \int x^\alpha pr(x)dx$
- Look for sequence of moments $[y_\alpha, \alpha = 0, 1, \ldots, \infty]$ to minimize the $\mathrm{E}[p(x)]$
- Relation between moments and probability distributions : LMI in terms of y

Infinite dimensional SDP

$$\underset{\text{vector } y}{\text{minimize}} \sum_\alpha p_\alpha y_\alpha \longrightarrow \text{Linear objective in terms of moments}$$

$$\text{subject to LMI}(y) \longrightarrow y \text{ is moment vector of probability distribution}$$

Step 3

(4) Optimization in terms of Moments

- **Decision variable:** $\text{moments: } y_\alpha = \mathrm{E}[x^\alpha] = \int x^\alpha pr(x)dx, \ \ \alpha = 0, \ldots, 2d$
- Look for sequence of moments $[y_\alpha, \alpha = 0, \ldots, 2d]$ to minimize the $\mathrm{E}[p(x)]$
- Truncate Matrices of SDP

SDP

$$\underset{y_\alpha, \alpha=0,\ldots,2d}{\text{minimize}} \sum_\alpha p_\alpha y_\alpha \longrightarrow \text{Linear objective in terms of moments}$$

$$\text{subject to LMI}_d(y)$$

4 Optimization in terms of Moments

- **Decision variable:** moments: $y_\alpha = \mathrm{E}[x^\alpha] = \int x^\alpha pr(x)dx, \ \alpha = 0, \dots, 2d$
- Look for sequence of moments $[y_\alpha, \alpha = 0, \dots, 2d]$ to minimize the $\mathrm{E}[p(x)]$
- Truncate Matrices of SDP

SDP

$$\begin{aligned} &\underset{y_\alpha, \alpha=0,\dots,2d}{\text{minimize}} \sum_\alpha p_\alpha y_\alpha \\ &\text{subject to } \mathrm{LMI}_d(y) \end{aligned}$$

→ Linear objective in terms of moments

- As we increase the number of the moments ($d \rightarrow \infty$), optimal value of the finite SDP converges to the optimal value of the original optimization.

4 Optimization in terms of Moments

- **Decision variable:** moments: $y_\alpha = \mathrm{E}[x^\alpha] = \int x^\alpha pr(x)dx, \;\; \alpha = 0, \ldots, 2d$
- Look for sequence of moments $[y_\alpha, \alpha = 0, \ldots, 2d]$ to minimize the $\mathrm{E}[p(x)]$
- Truncate Matrices of SDP

SDP

$$\underset{y_\alpha, \alpha=0,\ldots,2d}{\text{minimize}} \sum_\alpha p_\alpha y_\alpha \longrightarrow \text{Linear objective in terms of moments}$$

$$\text{subject to } \mathrm{LMI}_d(y)$$

- As we increase the number of the moments ($d \rightarrow \infty$), optimal value of the finite SDP converges to the optimal value of the original optimization.

- Extract the optimal solution of the original optimization, $\boldsymbol{x}^*$, from the optimal solution of the finite SDP, $[y^*_\alpha, \alpha = 0, \ldots, 2d]$

1. Nonlinear Optimization:

- Unconstrained Optimization $\Omega = \mathbb{R}^n$
- Constrained Optimization $\Omega = \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ i = 1, ..., m\}$

$$\mathrm{P}^* = \underset{x \in \Omega}{\text{minimize}} \quad p(x)$$

Nonlinear Optimization

2. Optimization in terms of Probability Distributions:

$$\underset{\substack{pr(x) \in \text{space of probability distributions} \\ \text{defined on } \Omega}}{\text{minimize}} \quad \mathrm{E}[p(x)] = \int p(x) pr(x) dx$$

Infinite dimensional LP

3. Optimization in terms of Moments

$$\underset{\text{vector } y}{\text{minimize}} \sum_{\alpha} p_\alpha y_\alpha$$
$$\text{subject to } \mathrm{LMI}(y)$$

Infinite dimensional SDP

4. Optimization in terms of Truncated Moments

$$\underset{y_\alpha, \alpha = 0, \ldots, 2d}{\text{minimize}} \sum_{\alpha} p_\alpha y_\alpha$$
$$\text{subject to } \mathrm{LMI}_d(y)$$

SDP

1. Nonlinear Optimization

2. Optimization in terms of Probability Distributions

Infinite dimensional LP

3. Optimization in terms of Moments

Infinite dimensional SDP

4. Optimization in terms of Truncated Moments

SDP

Extract the optimal solution x^*

1. Nonlinear Optimization

2. Optimization in terms of Probability Distributions

Infinite dimensional LP

Theory of Measure and Moments

3. Optimization in terms of Moments

Infinite dimensional SDP

4. Optimization in terms of Truncated Moments

SDP

Extract the optimal solution x^*

Topics

1) Overview of Measure and Moment based Optimization

2) Theory of Measure and Moments

3) Measure and Moment based Optimization

4) Examples and GloptiPoly Package

5) Proofs (Moment Conditions)

Theory of Measure and Moments

- Probability distributions and Measures
- Moments
- Representing Measure of Moments

Probability Space

Probability Space: (Ω, Σ, μ)

1) Ω : Set of all possible outcomes, (Sample Space)

2) $\sum$: Set of events (Borel $\sigma - algebra$ of Ω)
   - *Event:* a set of containing zero or more outcomes.

3) **Probability measure:** $\mu : \Sigma \rightarrow [0\ 1]$

For more information on "Borel $\sigma - algebra$" and "Borel measure": P. Cannarsa and T. D'Aprile "Lecture Notes on Measure Theory and Functional Analysis"

Probability Space

Probability Space: (Ω, Σ, μ)

1) Ω : Set of all possible outcomes, (Sample Space)

2) $\sum$: Set of events (Borel $\sigma - algebra$ of Ω)

   - *Event:* a set of containing zero or more outcomes.

3) **Probability measure:** $\mu : \Sigma \to [0\ 1]$

   Real-valued Function defined on a space of events $\sum$. (Finite (nonnegative) Borel measure on $\sum$)

For more information on "Borel $\sigma - algebra$" and "Borel measure": P. Cannarsa and T. D'Aprile "Lecture Notes on Measure Theory and Functional Analysis"

Probability Space

Probability Space: (Ω, Σ, μ)

1) Ω : Set of all possible outcomes, (Sample Space)

2) $\sum$: Set of events (Borel $\sigma - algebra$ of Ω)
   - *Event:* a set of containing zero or more outcomes.

3) **Probability measure:** $\mu : \Sigma \rightarrow [0\ \ 1]$

   Real-valued Function defined on a space of events $\sum$. (Finite (nonnegative) Borel measure on $\sum$)

Example: Given a random variable x with probability density function (pdf) $pr(x)$:

Probability that $x \in A$ $\quad \mu(A) = \underbrace{\int_A pr(x)dx}_{\text{Probability Measure}} \in [0\ \ 1]$

pdf

Probability Space

Probability Space: (Ω, Σ, μ)

1) Ω : Set of all possible outcomes, (Sample Space)

2) $\sum$: Set of events (Borel $\sigma - algebra$ of Ω)
    - *Event:* a set of containing zero or more outcomes.

3) **Probability measure:** $\mu : \Sigma \rightarrow [0\ 1]$
    Real-valued Function defined on a space of events $\sum$. (Finite (nonnegative) Borel measure on $\sum$)

Example: Given a random variable x with probability density function (pdf) $pr(x)$:

Probability that $x \in A$ $\quad \mu(A) = \underbrace{\int_A pr(x)dx}_{\text{Probability Measure}} \in [0\ 1]$

pdf

Set → Probability Measure → $[0\ 1]$

- Probability measure, measures the size of the set with respect to pdf.

Nonnegative Measure

- In general (nonnegative) measure μ assigns nonnegative real numbers to sets.

$\mu : \Sigma \to \mathbb{R}_+$ **(Nonnegative[1]) Measure**

$\mu : \Sigma \to [0\ 1]$ **Probability measure**

1: (Nonnegative) measure $\mu : \Sigma \to \mathbb{R}_+$, Signed Measures $\mu : \Sigma \to \mathbb{R}$

Nonnegative Measure

➢ In general (nonnegative) measure μ assigns nonnegative real numbers to sets.

$\mu : \Sigma \rightarrow \mathbb{R}_+$ **(Nonnegative) Measure**

$\mu : \Sigma \rightarrow [0\ 1]$ **Probability measure**

Example: Lebesgue measure

$$\mu_{leb}(A) = \int_A dx$$

- It measure the size of set A with respect to density function 1.
- **Examples:**
  - ➢ Lebesgue measure on $\mathbb{R}$

    $\mu_{leb}([-1,1]) = \int_{[-1,1]} dx = 2$ (Length)
  - ➢ Lebesgue measure on $\mathbb{R}^2$

    $\mu_{leb}([-1,1]^2) = \int\int_{[-1,1]^2} dx_1 dx_2 = 4$ (Area)
  - ➢ Lebesgue measure on $\mathbb{R}^3$

    $\mu_{leb}([-1,1]^3) = \int\int\int_{[-1,1]^3} dx_1 dx_2 dx_3 = 8$ (Volume)

Probability measure

$$\mu(A) = \underbrace{\int_A pr(x)dx}_{\text{Probability Measure}}$$

- It measure the size of set A with respect to probability density function.
- Probability of $x \in A$

- pdf satisfies $\int pr(x)dx = 1$

Nonnegative Measure

(Nonnegative) measure $\mu : \Sigma \to \mathbb{R}_+$

$$\mu(A) = \underbrace{\int_A d\mu}_{\text{Nonnegative Measure}} = \int_A f(x)dx$$

density function of μ

Nonnegative Measure

(Nonnegative) measure $\mu : \Sigma \to \mathbb{R}_+$

$$\mu(A) = \underbrace{\int_A d\mu}_{\text{Nonnegative Measure}} = \int_A f(x)dx$$

density function of μ

Parameters of a Measure 1) Density Function 2) Support Set 3) Moments

Nonnegative Measure

(Nonnegative) measure $\mu : \Sigma \rightarrow \mathbb{R}_+$

$$\mu(A) = \underbrace{\int_A d\mu}_{\text{Nonnegative Measure}} = \int_A f(x)dx$$

density function of μ

Parameters of a Measure 1) Density Function 2) Support Set 3) Moments

Support of a measure (*supp*)

A measure μ is said to be concentrated on a set "***S***" if $\mu(S^c) = 0$

complement set

- Smallest Closed Set S: support of μ

Nonnegative Measure

(Nonnegative) measure $\mu : \Sigma \to \mathbb{R}_+$

$$\mu(A) = \underbrace{\int_A d\mu}_{\text{Nonnegative Measure}} = \int_A f(x)dx$$

density function of μ

Parameters of a Measure 1) Density Function 2) Support Set 3) Moments

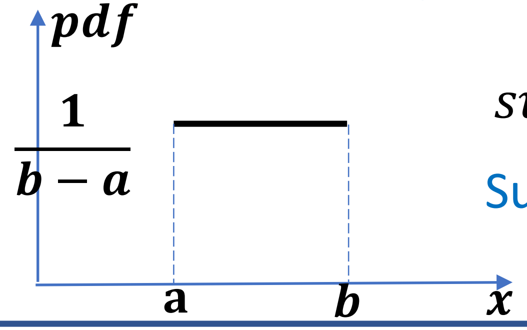
Support of a measure (*supp*)

A measure μ is said to be concentrated on a set "S" if $\mu(S^c) = 0$

complement set

➢ Smallest Closed Set S: support of μ

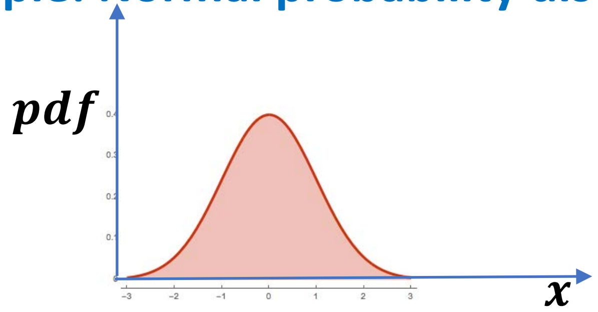
Example: Uniform probability distribution



$supp(U[a\ b]) = [\text{a b}]$

Support of uniform probability distribution on $[a, b]$

Example: Normal probability distribution



- Support of a probability measure is called **uncertainty set**.

Nonnegative Measure

(Nonnegative) measure $\mu : \Sigma \rightarrow \mathbb{R}_+$

$$\mu(A) = \underbrace{\int_A d\mu}_{\text{Nonnegative Measure}} = \int_A f(x)dx$$

density function of μ

Parameters of a Measure 1) Density Function 2) Support Set 3) Moments

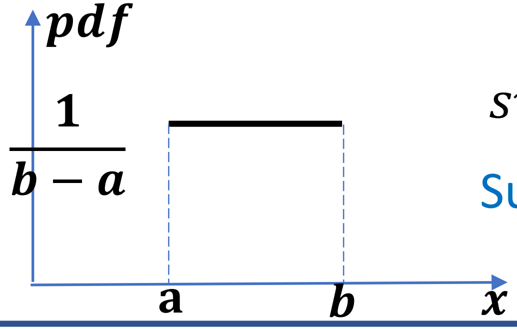
Support of a measure (*supp*)

A measure μ is said to be concentrated on a set "S" if $\mu(S^c) = 0$

complement set

➢ Smallest Closed Set S: support of μ

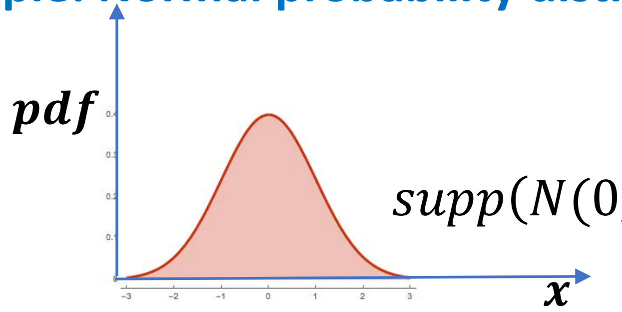
Example: Uniform probability distribution



$supp(U[a\ b]) = [\text{a b}]$

Support of uniform probability distribution on $[a, b]$

Example: Normal probability distribution



$supp(N(0,1)) = \mathbb{R}$

- Support of a **probability measure** is called **uncertainty set**.

1. Nonlinear Optimization

$$\mathrm{P}^* = \underset{x \in \Omega}{\text{minimize}} \quad p(x)$$

- Unconstrained Optimization $\Omega = \mathbb{R}^n$
- Constrained Optimization $\Omega = \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0, \ i = 1, ..., m\}$

2. Optimization in Measure Space

$$\mathrm{P}^*_\mu = \underset{\mu \in \mathcal{M}(\Omega)}{\text{minimize}} \ \ \mathrm{E}_\mu[p(x)] = \int p(x) d\mu$$

$$\text{subject to} \ \int_\Omega d\mu = 1$$

$\longrightarrow$ Probability Measure

$\mathcal{M}(\Omega)$: space of measures supported on Ω

Moments of a Measure

Given $(\alpha_1, \ldots, \alpha_n), \;\; \alpha_i \in \mathbb{N}, \;\; \alpha = \sum_{i=1}^{n} \alpha_i$

- **moment** of order α of a measure μ in $\mathbb{R}^n$ is defined as

$$y_\alpha = \mathrm{E}_\mu[x^\alpha] = \int x^\alpha d\mu = \int x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n} d\mu$$

Moments of a Measure

Given $(\alpha_1, ..., \alpha_n), \ \alpha_i \in \mathbb{N}, \ \alpha = \sum_{i=1}^{n} \alpha_i$

- **moment** of order α of a measure μ in $\mathbb{R}^n$ is defined as

$$y_\alpha = \mathrm{E}_\mu[x^\alpha] = \int x^\alpha d\mu = \int x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n} d\mu$$

Example: **Measure μ in $\mathbb{R}$**

- 1-st moment (mean): $\longrightarrow$ $y_1 = \mathrm{E}_\mu[x^1] = \int x d\mu$
- 2-nd moment: $\longrightarrow$ $y_2 = \mathrm{E}_\mu[x^2] = \int x^2 d\mu$
- 0-moment: (volume of measure): $y_0 = \mathrm{E}_\mu[x^0] = \int d\mu$

Moments of a Measure

Given $(\alpha_1, \ldots, \alpha_n), \ \alpha_i \in \mathbb{N}, \ \ \alpha = \sum_{i=1}^{n} \alpha_i$

- **moment** of order α of a measure μ in $\mathbb{R}^n$ is defined as

$$y_\alpha = \mathrm{E}_\mu[x^\alpha] = \int x^\alpha d\mu = \int x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n} d\mu$$

Example: **Measure μ in $\mathbb{R}$**

- 1-st moment (mean): $\longrightarrow$ $y_1 = \mathrm{E}_\mu[x^1] = \int x d\mu$
- 2-nd moment: $\longrightarrow$ $y_2 = \mathrm{E}_\mu[x^2] = \int x^2 d\mu$

$$var = \mathrm{E}[(x - \mathrm{E}_\mu[x])^2] = y_2 - y_1^2$$

- 0-moment: (volume of measure): $y_0 = \mathrm{E}_\mu[x^0] = \int d\mu$

Moments of a Measure

Given $(\alpha_1, ..., \alpha_n), \;\; \alpha_i \in \mathbb{N}, \;\; \alpha = \sum_{i=1}^{n} \alpha_i$

- **moment** of order α of a measure μ in $\mathbb{R}^n$ is defined as

$$y_\alpha = \mathrm{E}_\mu[x^\alpha] = \int x^\alpha d\mu = \int x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n} d\mu$$

Example: **Measure μ in $\mathbb{R}$**

- 1-st moment (mean): $\longrightarrow$ $y_1 = \mathrm{E}_\mu[x^1] = \int x d\mu$
- 2-nd moment: $\longrightarrow$ $y_2 = \mathrm{E}_\mu[x^2] = \int x^2 d\mu$

$$var = \mathrm{E}[(x - \mathrm{E}_\mu[x])^2] = y_2 - y_1^2$$

- 0-moment: (volume of measure): $y_0 = \mathrm{E}_\mu[x^0] = \int d\mu$

- Probability measure μ: $y_0 = \mathrm{E}_\mu[x^0] = \int d\mu = \int pr(x) dx = 1$

Moments of a Measure

Given $(\alpha_1, ..., \alpha_n), \; \alpha_i \in \mathbb{N}, \;\; \alpha = \sum_{i=1}^{n} \alpha_i$

- **moment** of order α of a measure μ in $\mathbb{R}^n$ is defined as

$$y_\alpha = \mathrm{E}_\mu[x^\alpha] = \int x^\alpha d\mu = \int x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n} d\mu$$

Example: **Measure μ in $\mathbb{R}^2$**

Moments of a Measure

Given $(\alpha_1, ..., \alpha_n), \ \alpha_i \in \mathbb{N}, \ \ \alpha = \sum_{i=1}^{n} \alpha_i$

- ➢ **moment** of order α of a measure μ in $\mathbb{R}^n$ is defined as

$$y_\alpha = \mathrm{E}_\mu[x^\alpha] = \int x^\alpha d\mu = \int x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n} d\mu$$

Example: **Measure μ in $\mathbb{R}^2$**

- 1-st moments (mean): → $y_{10} = \mathrm{E}[x_1] = \int x_1 d\mu$ $\quad y_{01} = \mathrm{E}[x_2] = \int x_2 d\mu$

Moments of a Measure

Given $(\alpha_1, ..., \alpha_n), \ \alpha_i \in \mathbb{N}, \ \ \alpha = \sum_{i=1}^{n} \alpha_i$

- **moment** of order α of a measure μ in $\mathbb{R}^n$ is defined as

$$y_\alpha = \mathrm{E}_\mu[x^\alpha] = \int x^\alpha d\mu = \int x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n} d\mu$$

Example: **Measure μ in $\mathbb{R}^2$**

- 1-st moments (mean): → $y_{10} = \mathrm{E}[x_1] = \int x_1 d\mu$ $\quad y_{01} = \mathrm{E}[x_2] = \int x_2 d\mu$
- 2-nd moments:

$$y_{20} = \mathrm{E}[x_1^2] = \int x_1^2 d\mu \qquad y_{11} = \mathrm{E}[x_1 x_2] = \int x_1 x_2 d\mu \qquad y_{02} = \mathrm{E}[x_2^2] = \int x_2^2 d\mu$$

- α order moments: → Vector of Expected values of all the monomials of order α

Moments of a Measure

Given $(\alpha_1, ..., \alpha_n)$, $\alpha_i \in \mathbb{N}$, $\alpha = \sum_{i=1}^{n} \alpha_i$

- **moment** of order α of a measure μ in $\mathbb{R}^n$ is defined as

$$y_\alpha = \mathrm{E}_\mu[x^\alpha] = \int x^\alpha d\mu = \int x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n} d\mu$$

Example: **Measure μ in $\mathbb{R}^2$**

- 1-st moments (mean): → $y_{10} = \mathrm{E}[x_1] = \int x_1 d\mu$ $\quad y_{01} = \mathrm{E}[x_2] = \int x_2 d\mu$
- 2-nd moments:

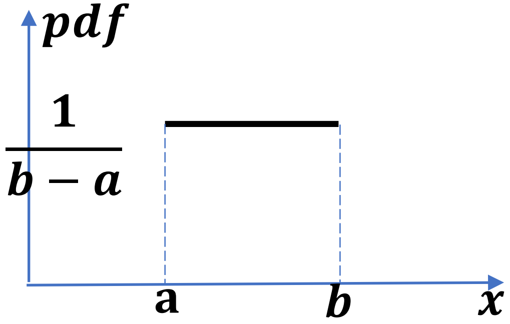
$$y_{20} = \mathrm{E}[x_1^2] = \int x_1^2 d\mu \qquad y_{11} = \mathrm{E}[x_1 x_2] = \int x_1 x_2 d\mu \qquad y_{02} = \mathrm{E}[x_2^2] = \int x_2^2 d\mu$$

- α order moments: → Vector of Expected values of all the monomials of order α

- Number of moments of a measure in $\mathbb{R}^n$ up to order d: $\quad S_{n,d} = \binom{n+d}{n} = \frac{(n+d)!}{n!d!}$

Example: Probability Distributions and Moments

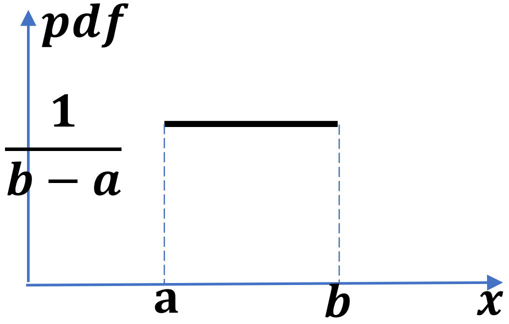
Uniform Probability Distribution $\mathrm{U}[a, b]$



$$pr(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{otherwise} \end{cases}$$

Example: Probability Distributions and Moments

Uniform Probability Distribution $\mathrm{U}[a, b]$



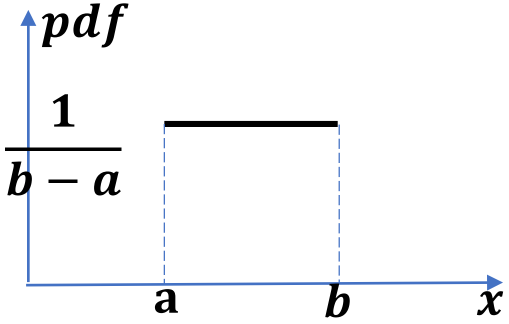
$$pr(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{otherwise} \end{cases}$$

$$y_\alpha = \mathrm{E}[x^\alpha] = \int x^\alpha d\mu = \int x^\alpha pr(x) dx = \int_a^b x^\alpha \frac{1}{b-a} dx = \frac{1}{b-a} \frac{x^{\alpha+1}}{\alpha+1} \Big|_{x=a}^{x=b}$$

$$y_\alpha = \frac{1}{b-a} \frac{b^{\alpha+1} - a^{\alpha+1}}{\alpha+1}$$

Example: Probability Distributions and Moments

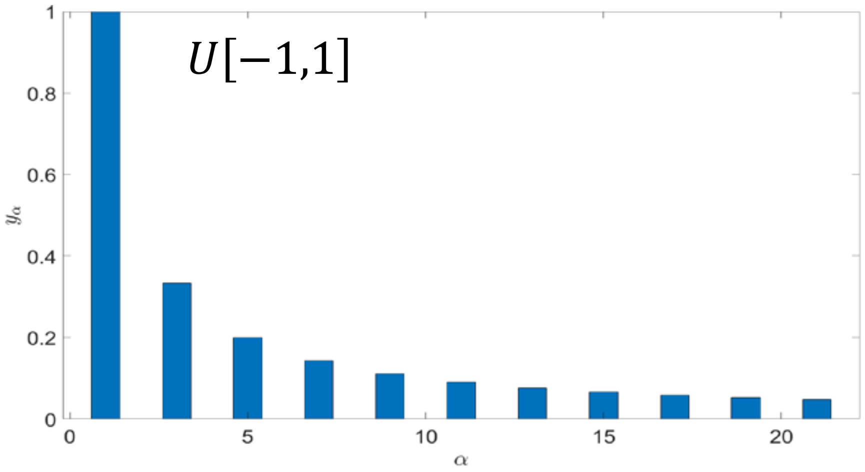
Uniform Probability Distribution U$[a, b]$

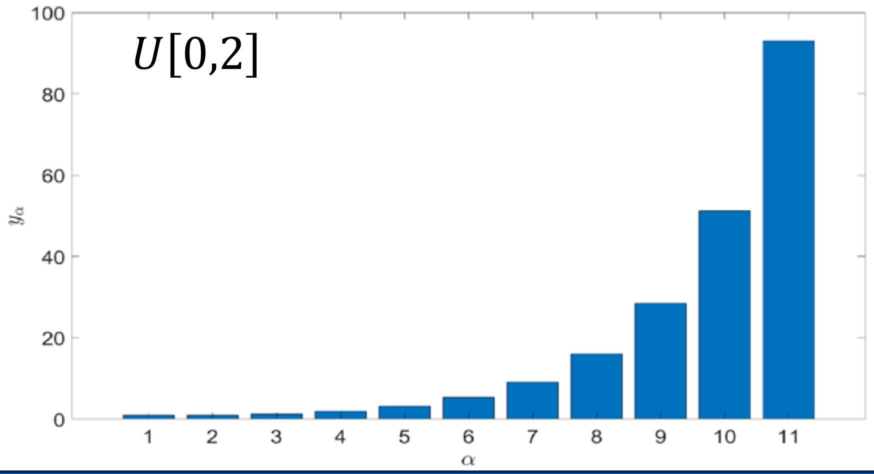


$$pr(x) = \begin{cases} \frac{1}{b-a} & x \in [a,b] \\ 0 & \text{otherwise} \end{cases}$$

$$y_\alpha = \mathrm{E}[x^\alpha] = \int x^\alpha d\mu = \int x^\alpha pr(x)dx = \int_a^b x^\alpha \frac{1}{b-a}dx = \frac{1}{b-a}\frac{x^{\alpha+1}}{\alpha+1}\Big|_{x=a}^{x=b}$$

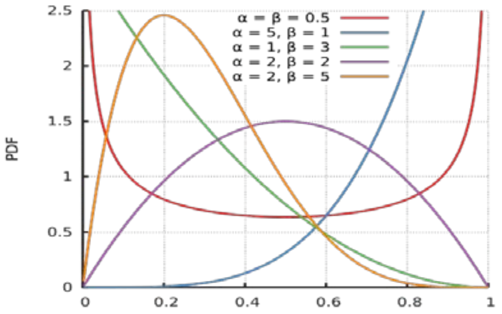
$$y_\alpha = \frac{1}{b-a}\frac{b^{\alpha+1}-a^{\alpha+1}}{\alpha+1}$$





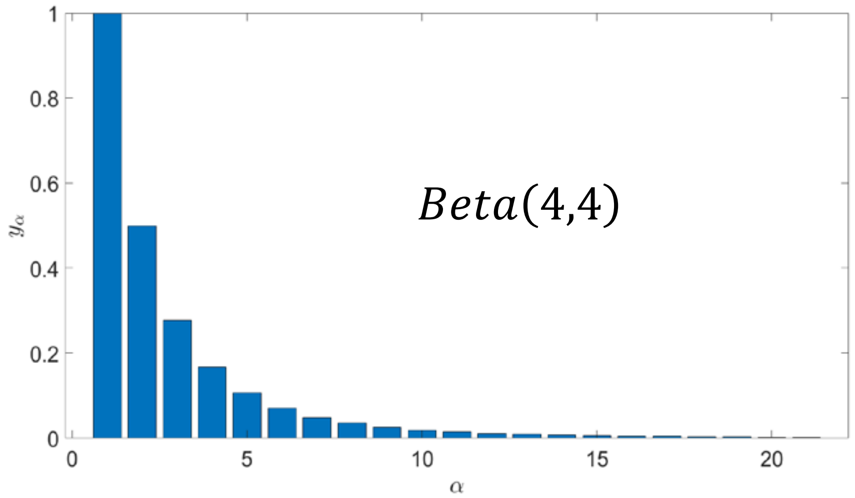
Example: Probability Distributions and Moments

Beta Distribution $Beta(a, b)$



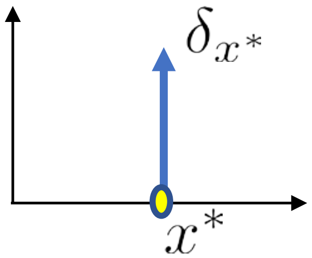
pdf $pr(x) = \dfrac{x^{a-1}(1-x)x^{b-1}}{B(a,b)}$

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$



$$y_\alpha = \frac{a+\alpha-1}{a+b+\alpha-1} y_{\alpha-1}, \quad y_0 = 1$$

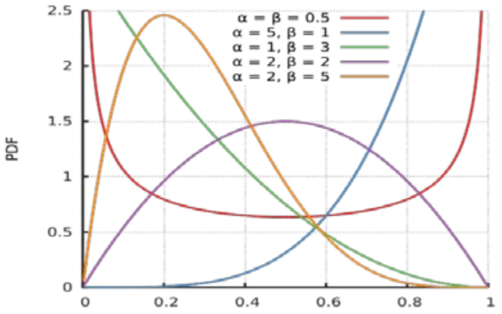
Dirac distribution $\boldsymbol{\delta_{x^*}}$



$$y_\alpha = E[x^\alpha] = \int x^\alpha \delta_{x^*} dx = {x^*}^\alpha$$

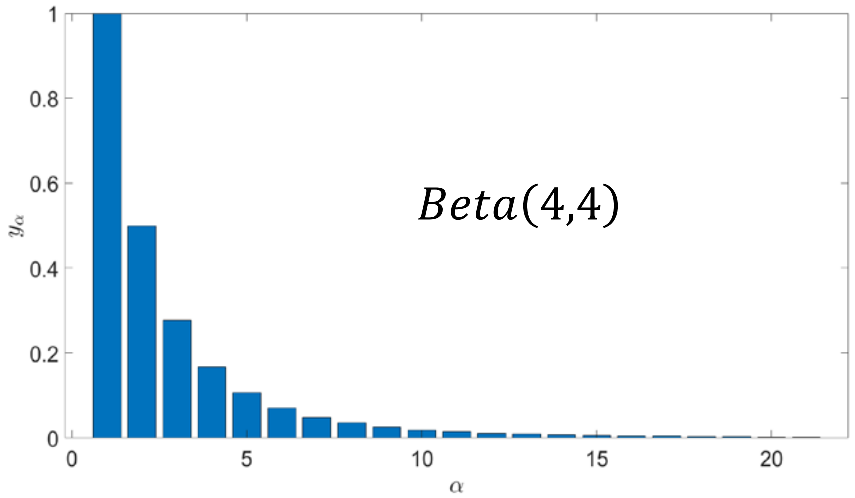
Example: Probability Distributions and Moments

Beta Distribution $Beta(a, b)$



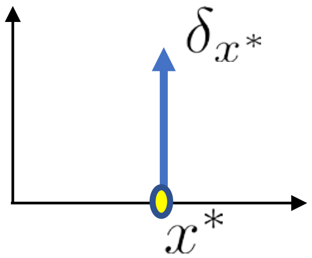
pdf $pr(x) = \dfrac{x^{a-1}(1-x)x^{b-1}}{B(a,b)}$

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$



$$y_\alpha = \frac{a + \alpha - 1}{a + b + \alpha - 1} y_{\alpha-1}, \quad y_0 = 1$$

Dirac distribution $\boldsymbol{\delta_{x^*}}$



$$y_\alpha = E[x^\alpha] = \int x^\alpha \delta_{x^*} dx = {x^*}^\alpha \qquad y_1 = E[x^1] = x^*$$

- **By looking at the first moment of a Dirac measure, we can extract the optimal solution x^***

Representing Measure of Moments

1. Nonlinear Optimization

2. Optimization in terms of Probability Distributions

Infinite dimensional LP

Relation between Measure and Moments

3. Optimization in terms of Moments

Infinite dimensional SDP

4. Optimization in terms of Truncated Moments

SDP

Extract the optimal solution x^*

Representing Measure

- Instead of working with measure, we want to work with its moment sequence.

Representing Measure

- Instead of working with measure, we want to work with its moment sequence.
- A measure μ is said to be a **representing measure** of a real sequence $y = [y_\alpha, \alpha = 0, ..., d]$ If

Representing Measure

- Instead of working with measure, we want to work with its moment sequence.
- A measure μ is said to be a **representing measure** of a real sequence $y = [y_\alpha, \alpha = 0, ..., d]$
  If sequence y is the moments sequence of the measure μ , i.e.,$y_\alpha = E_\mu[x^\alpha]$

Representing Measure

- Instead of working with measure, we want to work with its moment sequence.

- A measure μ is said to be a **representing measure** of a real sequence $y = [y_\alpha, \alpha = 0, ..., d]$ If sequence y is the moments sequence of the measure μ , i.e., $y_\alpha = E_\mu[x^\alpha]$

> - Measure μ represents the sequence $y = [y_\alpha, \alpha = 0, ..., d]$ in the measure space.
> - Sequence $y = [y_\alpha, \alpha = 0, ..., d]$ represents measure μ in the moment space.

Representing Measure

- Instead of working with measure, we want to work with its moment sequence.

- A measure μ is said to be a **representing measure** of a real sequence $y = [y_\alpha, \alpha = 0, ..., d]$
  If sequence y is the moments sequence of the measure μ , i.e., $y_\alpha = E_\mu[x^\alpha]$

> - Measure μ represents the sequence $y = [y_\alpha, \alpha = 0, ..., d]$ in the measure space.
> - Sequence $y = [y_\alpha, \alpha = 0, ..., d]$ represents measure μ in the moment space.

1) Can we uniquely represent a measure by the (infinite) sequence of its moments?

Representing Measure

- Instead of working with measure, we want to work with its moment sequence.
- A measure μ is said to be a **representing measure** of a real sequence $y = [y_\alpha, \alpha = 0, ..., d]$
  If sequence y is the moments sequence of the measure μ , i.e.,$y_\alpha = E_\mu[x^\alpha]$

➢ Measure μ represents the sequence $y = [y_\alpha, \alpha = 0, ..., d]$ in the measure space.

➢ Sequence $y = [y_\alpha, \alpha = 0, ..., d]$ represents measure μ in the moment space.

1) Can we uniquely represent a measure by the (infinite) sequence of its moments?

Can we find measures with the same moments sequence ?

Representing Measure

- Instead of working with measure, we want to work with its moment sequence.

- A measure μ is said to be a **representing measure** of a real sequence $y = [y_\alpha, \alpha = 0, ..., d]$ If sequence y is the moments sequence of the measure μ , i.e., $y_\alpha = E_\mu[x^\alpha]$

➢ Measure μ represents the sequence $y = [y_\alpha, \alpha = 0, ..., d]$ in the measure space.

➢ Sequence $y = [y_\alpha, \alpha = 0, ..., d]$ represents measure μ in the moment space.

1) Can we uniquely represent a measure by the (infinite) sequence of its moments?

Can we find measures with the same moments sequence ?

2) Does every sequence y have a *representing measure* μ ?

Representing Measure

1) Can we uniquely represent a measure by the (infinite) sequence of its moments?

Can we find measures with the same moments sequence ?

- *Chapter 3*: Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

Representing Measure

1) Can we uniquely represent a measure by the (infinite) sequence of its moments?

Can we find measures with the same moments sequence ?

Example:

The following 2 measure have the same moment sequence:

$$pr(x) = \begin{cases} \frac{1}{x\sqrt{2\pi}} \exp -\frac{\ln(x)^2}{2} & \text{if } x > 0 \\ 0 \quad \text{otherwise} \end{cases}$$

pdf of "Log-Normal" distribution

$$pr_a(x) = pr(x)[1 + a\sin(2\pi \ln(\mathrm{x}))]$$

- *Chapter 3*: Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

Representing Measure

1) Can we uniquely represent a measure by the (infinite) sequence of its moments?

Can we find measures with the same moments sequence ?

➢ A measure supported on a **compact set** is uniquely determined by the (infinite) sequence of its moments.

- *Chapter 3*: Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

Representing Measure

1) Can we uniquely represent a measure by the (infinite) sequence of its moments?

Can we find measures with the same moments sequence ?

➢ A measure supported on a **compact set** is uniquely determined by the (infinite) sequence of its moments.

- In this case **representing measure** of the moment sequence is unique.
- This unique Measure is said to be **determinate** (determined by its moments)

- *Chapter 3*: Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

Representing Measure

2) Does every sequence y have a *representing measure* μ ?

Representing Measure

2) Does every sequence y have a *representing measure* μ ?

Example

Does sequence $y = [1, 0.5, 0.2]$ have a *representing measure* ?

$$\overbrace{\mu,\ y_0 = 1,\ y_1 = 0.5,\ y_2 = 0.2}$$

Representing Measure

2) Does every sequence y have a *representing measure* μ ?

Example

Does sequence $y = [1, 0.5, 0.2]$ have a *representing measure* ?

$$\overbrace{\mu, \;\; y_0 = 1, \;\; y_1 = 0.5, \;\; y_2 = 0.2}$$

$$var = \mathrm{E}[(x - \mathrm{E}[x])^2] \geq 0 \Longrightarrow var = \mathrm{E}[(x - \mathrm{E}[x])^2] = y_2 - y_1^2 \geq 0$$

Representing Measure

2) Does every sequence y have a *representing measure* μ ?

Example

Does sequence $y = [1, 0.5, 0.2]$ have a *representing measure* ?

$$\overbrace{\mu,\ y_0 = 1,\ y_1 = 0.5,\ y_2 = 0.2}$$

$$var = \mathrm{E}[(x - \mathrm{E}[x])^2] \geq 0 \Rightarrow var = \mathrm{E}[(x - \mathrm{E}[x])^2] = y_2 - y_1^2 \geq 0$$

$y = [1, 0.5, 0.2]$ does not satisfy this condition; Hence it does not have a representing measure.

Representing Measure

2) Does every sequence y have a *representing measure* μ ?

- Given an arbitrary sequence y, there might not be any representing measure μ (all of whose moments coincide with the y).

- Not every sequence y has a *representing measure* μ.

- **What are the constraints for sequence y to have a representing measures ?**

Representing Measure

2) Does every sequence y have a *representing measure* μ ?

- Given an arbitrary sequence y, there might not be any representing measure μ (all of whose moments coincide with the y).

- Not every sequence y has a *representing measure* μ.

- **What are the constraints for sequence y to have a representing measures ?**

Example: LMI Condition on the moments up to order 2 (probability measure)

$$var = \mathrm{E}[(x - \mathrm{E}[x])^2] = y_2 - y_1^2 \geq 0 \Longrightarrow \begin{bmatrix} 1 & y_1 \\ y_1 & y_2 \end{bmatrix} \succcurlyeq 0$$

Representing Measure

2) Does every sequence y have a *representing measure* μ ?

- Given an arbitrary sequence y, there might not be any representing measure μ (all of whose moments coincide with the y).

- Not every sequence y has a *representing measure* μ.

- **What are the constraints for sequence y to have a representing measures ?**

Example: LMI Condition on the moments up to order 2 (probability measure)

$$var = \mathrm{E}[(x - \mathrm{E}[x])^2] = y_2 - y_1^2 \geq 0 \quad \Longrightarrow \quad \begin{bmatrix} 1 & y_1 \\ y_1 & y_2 \end{bmatrix} \succcurlyeq 0$$

- **Moments of a measure should satisfy Linear Matrix Inequalities in terms of**

i) Moment Matrix ii) Localizing Matrix

Representing Measure

- Moments of a measure should satisfy Linear Matrix Inequalities in terms of
  i) Moment Matrix **ii) localizing Matrix**

- **Moment Matrix** to characterize moments of measure μ in $\mathbb{R}^n$
- **Moment and Localizing Matrix** to characterize moments of measure μ on compact set $\mathbf{K}$

1. Nonlinear Optimization

$$\text{P}^* = \underset{x \in \Omega}{\text{minimize}} \quad p(x)$$

- Unconstrained Optimization $\Omega = \mathbb{R}^n$
- Constrained Optimization $\Omega = \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0, \ i = 1, ..., m\}$

2. Optimization in Measure Space

$$\text{P}_\mu^* = \underset{\mu \in \mathcal{M}(\Omega)}{\text{minimize}} \ \ \text{E}_\mu[p(x)] = \int p(x) d\mu$$

$$\text{subject to} \ \int_\Omega d\mu = 1 \longrightarrow \text{Probability Measure}$$

$\mathcal{M}(\Omega)$: space of measures supported on Ω

3. Optimization in Moment Space

$$\text{P}_{mom}^* = \underset{y}{\text{minimize}} \ \ \text{E}_\mu[p(x)] = \int p(x) d\mu = \sum_\alpha p_\alpha y_\alpha$$

$$\text{subject to} \ y_0 = 1 \longrightarrow \text{Probability Measure}$$

y is moment vector of probability distribution $\longleftarrow$ $\text{Moment Matrix } (y) \succcurlyeq 0 \quad \text{Localizing Matrix } (g_i y) \succcurlyeq 0$

Moment Matrix

Moment Matrix associated with a sequence of moments up to order $2d$ is the real symmetric square matrix:

Moment Matrix of order d: $$\mathbf{M}_d(y) = \mathrm{E}_\mu[B_d(x)B_d^T(x)]$$

- $B_d\,(x)$: vector of monomial up to order d

Moment Matrix

Moment Matrix associated with a sequence of moments up to order $2d$ is the real symmetric square matrix:

Moment Matrix of order d: $$\mathbf{M}_d(y) = \mathrm{E}_\mu[B_d(x)B_d^T(x)]$$

- $B_d\,(x)$: vector of monomial up to order d

Example: Moment matrix of order $d = 2$ of a measure in $\mathbb{R}$

$$B_2(x) = \begin{bmatrix} 1 \\ x_1 \\ x_1^2 \end{bmatrix} \qquad \mathbf{M}_2(y) = \mathrm{E}[B_2(x)B_2^T(x)] = \mathrm{E}_\mu[\begin{bmatrix} 1 \\ x_1 \\ x_1^2 \end{bmatrix} \begin{bmatrix} 1 & x_1 & x_1^2 \end{bmatrix}] = \mathrm{E}_\mu[\begin{bmatrix} 1 & x_1 & x_1^2 \\ x_1 & x_1^2 & x_1^3 \\ x_1^2 & x_1^3 & x_1^4 \end{bmatrix}] = \begin{bmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix}$$

vector of monomials up to order d=2

Moment Matrix

Moment Matrix associated with a sequence of moments up to order $2d$ is the real symmetric square matrix:

Moment Matrix of order d: $$\mathbf{M}_d(y) = \mathrm{E}_\mu[B_d(x)B_d^T(x)]$$

- $B_d\,(x)$: vector of monomial up to order d

Example: Moment matrix of order $d = 2$ of a measure in $\mathbb{R}$

$$B_2(x) = \begin{bmatrix} 1 \\ x_1 \\ x_1^2 \end{bmatrix} \qquad \mathbf{M}_2(y) = \mathrm{E}[B_2(x)B_2^T(x)] = \mathrm{E}_\mu[\begin{bmatrix} 1 \\ x_1 \\ x_1^2 \end{bmatrix}\begin{bmatrix} 1 & x_1 & x_1^2 \end{bmatrix}] = \mathrm{E}_\mu[\begin{bmatrix} 1 & x_1 & x_1^2 \\ x_1 & x_1^2 & x_1^3 \\ x_1^2 & x_1^3 & x_1^4 \end{bmatrix}] = \begin{bmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix}$$

vector of monomials up to order d=2

➢ $\mathbf{M}_d(y)$: requires the moments up to order 2d

Moment Matrix

Moment Matrix associated with a sequence of moments up to order $2d$ is the real symmetric square matrix:

Moment Matrix of order d: $$\mathbf{M}_d(y) = \mathrm{E}_\mu[B_d(x)B_d^T(x)]$$

- $B_d\,(x)$: vector of monomial up to order d

Example: Moment matrix of order $d = 2$ of a measure in $\mathbb{R}$

$$B_2(x) = \begin{bmatrix} 1 \\ x_1 \\ x_1^2 \end{bmatrix} \qquad \mathbf{M}_2(y) = \mathrm{E}[B_2(x)B_2^T(x)] = \mathrm{E}_\mu[\begin{bmatrix} 1 \\ x_1 \\ x_1^2 \end{bmatrix}\begin{bmatrix} 1 & x_1 & x_1^2 \end{bmatrix}] = \mathrm{E}_\mu[\begin{bmatrix} 1 & x_1 & x_1^2 \\ x_1 & x_1^2 & x_1^3 \\ x_1^2 & x_1^3 & x_1^4 \end{bmatrix}] = \begin{bmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix}$$

vector of monomials up to order d=2

- ➢ $\mathbf{M}_d(y)$: requires the moments up to order 2d
- ➢ Size of the moment matrix of order d: $S_{n,d} \times S_{n,d}: \quad S_{n,d} = \binom{n+d}{n} = \frac{(n+d)!}{n!d!}$
- ➢ Number of the moments $S_{n,2d}: \quad S_{n,2d} = \binom{n+2d}{n} = \frac{(n+2d)!}{n!2d!}$

Example:

Moment Matrix of moments up to order d=2 of a Measure in $\mathbb{R}^2$

$$\mathbf{M}_2(y) = \mathrm{E}_\mu[B_2(x)B_2^T(x)] = \mathrm{E}_\mu[\begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{bmatrix} [1 \quad x_1 \quad x_2 \quad x_1^2 \quad x_1x_2 \quad x_2^2]] =$$

$$= \mathrm{E}_\mu \begin{bmatrix} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 \\ x_1 & x_1^2 & x_1x_2 & x_1^3 & x_1^2x_2 & x_1x_2^2 \\ x_2 & x_2x_1 & x_2^2 & x_2x_1^2 & x_1x_2^2 & x_2^3 \\ x_1^2 & x_1^3 & x_1^2x_2 & x_1^4 & x_1^3x_2 & x_1^2x_2^2 \\ x_1x_2 & x_1^2x_2 & x_1x_2^2 & x_1^3x_2 & x_1^2x_2^2 & x_1x_2^3 \\ x_2^2 & x_2^2x_1 & x_2^3 & x_2^2x_1^2 & x_1x_2^3 & x_2^4 \end{bmatrix}] = \begin{bmatrix} y_{00} & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ y_{01} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{bmatrix}$$

Example:

Moment Matrix of moments up to order d=2 of a Measure in $\mathbb{R}^2$

$$\mathbf{M}_2(y) = \mathrm{E}_\mu[B_2(x)B_2^T(x)] = \mathrm{E}_\mu[\begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{bmatrix} [1 \quad x_1 \quad x_2 \quad x_1^2 \quad x_1x_2 \quad x_2^2]] =$$

$$= \mathrm{E}_\mu \left[\begin{bmatrix} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 \\ x_1 & x_1^2 & x_1x_2 & x_1^3 & x_1^2x_2 & x_1x_2^2 \\ x_2 & x_2x_1 & x_2^2 & x_2x_1^2 & x_1x_2^2 & x_2^3 \\ x_1^2 & x_1^3 & x_1^2x_2 & x_1^4 & x_1^3x_2 & x_1^2x_2^2 \\ x_1x_2 & x_1^2x_2 & x_1x_2^2 & x_1^3x_2 & x_1^2x_2^2 & x_1x_2^3 \\ x_2^2 & x_2^2x_1 & x_2^3 & x_2^2x_1^2 & x_1x_2^3 & x_2^4 \end{bmatrix} \right] = \begin{bmatrix} y_{00} & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ y_{10} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{bmatrix}$$

Moments of order 1

Moments of order 2

Moments of order 3

Moments of order 4

Example:

Moment Matrix of moments up to order d=2 of a Measure in $\mathbb{R}^2$

$$\mathbf{M}_2(y) = \mathrm{E}_\mu[B_2(x)B_2^T(x)] = \mathrm{E}_\mu[\begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{bmatrix} [1 \quad x_1 \quad x_2 \quad x_1^2 \quad x_1x_2 \quad x_2^2]] =$$

$$= \mathrm{E}_\mu \begin{bmatrix} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 \\ x_1 & x_1^2 & x_1x_2 & x_1^3 & x_1^2x_2 & x_1x_2^2 \\ x_2 & x_2x_1 & x_2^2 & x_2x_1^2 & x_1x_2^2 & x_2^3 \\ x_1^2 & x_1^3 & x_1^2x_2 & x_1^4 & x_1^3x_2 & x_1^2x_2^2 \\ x_1x_2 & x_1^2x_2 & x_1x_2^2 & x_1^3x_2 & x_1^2x_2^2 & x_1x_2^3 \\ x_2^2 & x_2^2x_1 & x_2^3 & x_2^2x_1^2 & x_1x_2^3 & x_2^4 \end{bmatrix}] = \begin{bmatrix} y_{00} & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ y_{10} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{bmatrix}$$

Moments of order 1

Moments of order 2

Moments of order 3

Moments of order 4

Moment Condition

Moments of every nonnegative measure μ in $\mathbb{R}^n$ satisfies :

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \qquad \text{(PSD)}$$

Moment Condition

Moments of every nonnegative measure μ in $\mathbb{R}^n$ satisfies :

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \qquad \text{(PSD)}$$

- Moment condition relies on the information of all the moments, i.e., $\mathbf{M}_\infty(y) \succcurlyeq 0$

Moment Condition

Moments of every nonnegative measure μ in $\mathbb{R}^n$ satisfies :

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \qquad \text{(PSD)}$$

- Moment condition relies on the information of all the moments, i.e., $\mathbf{M}_\infty(y) \succcurlyeq 0$
- Not every sequence y that satisfies $\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d$, has a representing measure μ in $\mathbb{R}^n$

$$\boxed{\mu\ ,\ y\text{:moments}} \;\underset{\not\Leftarrow}{\Rightarrow}\; \boxed{\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d}$$

Moment Condition

Moments of every nonnegative measure μ in $\mathbb{R}^n$ satisfies :

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \qquad \text{(PSD)}$$

- Moment condition relies on the information of all the moments, i.e., $\mathbf{M}_\infty(y) \succcurlyeq 0$
- Not every sequence y that satisfies $\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d$, has a representing measure μ in $\mathbb{R}^n$

$$\boxed{\mu \ , \ y\text{:moments}} \ \underset{\not\Leftarrow}{\Rightarrow} \ \boxed{\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d}$$

- It is similar to SOS condition. Not every nonnegative polynomial has a SOS representation.

Localizing Matrix

Localizing matrix associated with a sequence of moments up to order $2d$ and a polynomial $g(x) \in \mathbb{R}[x]$ is the real symmetric square matrix:

Localizing Matrix: $$\mathbf{M}_{d-d_g}(gy) = \mathrm{E}_{\mu}[g(x)B_{d-d_g}(x)B^T_{d-d_g}(x)]$$

- B_{d-d_g} : vector of monomial up to order $d - d_g$
- $d_g = \left\lceil \frac{\deg(g(x))}{2} \right\rceil$ the smallest integer larger than $\frac{\deg(g(x))}{2}$

Localizing Matrix

Localizing matrix associated with a sequence of moments up to order $2d$ and a polynomial $g(x) \in \mathbb{R}[x]$ is the real symmetric square matrix:

Localizing Matrix: $$\mathbf{M}_{d-d_g}(gy) = \mathrm{E}_{\mu}[g(x) B_{d-d_g}(x) B^T_{d-d_g}(x)]$$

- B_{d-d_g} : vector of monomial up to order $d - d_g$
- $d_g = \left\lceil \frac{\deg(g(x))}{2} \right\rceil$ the smallest integer larger than $\frac{\deg(g(x))}{2}$

➢ $\mathbf{M}_{d-d_g}(gy)$: requires the moments up to order 2d because $deg\left(g(x) B_{d-d_g}(x) B^T_{d-d_g}(x)\right) \leq 2d$

Localizing Matrix

Localizing matrix associated with a sequence of moments up to order $2d$ and a polynomial $g(x) \in \mathbb{R}[x]$ is the real symmetric square matrix:

Localizing Matrix: $$\mathbf{M}_{d-d_g}(gy) = \mathrm{E}_\mu[g(x)B_{d-d_g}(x)B^T_{d-d_g}(x)]$$

- B_{d-d_g} : vector of monomial up to order $d - d_g$
- $d_g = \left\lceil \frac{\deg(g(x))}{2} \right\rceil$ the smallest integer larger than $\frac{\deg(g(x))}{2}$

➢ $\mathbf{M}_{d-d_g}(gy)$: requires the moments up to order 2d because $deg\left(g(x)B_{d-d_g}(x)B^T_{d-d_g}(x)\right) \leq 2d$

$$\deg\left(g(x)B_{d-d_g}(x)B_{d-d_g}(x)\right) = 2(d-d_g) + \deg(g(x)) = 2d - 2\left\lceil \frac{\deg(g(x))}{2} \right\rceil + \deg(g(x)) = \begin{cases} 2d & \text{if } \deg(g(x))\text{:even} \\ 2d-1 & \text{if } \deg(g(x))\text{:odd} \end{cases}$$

Localizing Matrix

Localizing matrix associated with a sequence of moments up to order $2d$ and a polynomial $g(x) \in \mathbb{R}[x]$ is the real symmetric square matrix:

Localizing Matrix: $$\mathbf{M}_{d-d_g}(gy) = \mathrm{E}_\mu[g(x) B_{d-d_g}(x) B^T_{d-d_g}(x)]$$

- B_{d-d_g} : vector of monomial up to order $d - d_g$
- $d_g = \left\lceil \frac{\deg(g(x))}{2} \right\rceil$ the smallest integer larger than $\frac{\deg(g(x))}{2}$

Example: Localizing matrix for moments up to order 4 of a Measure μ in $\mathbb{R}^2$ and polynomial $g(x) = a - x_1^2 - x_2^2$

Localizing Matrix

Localizing matrix associated with a sequence of moments up to order $2d$ and a polynomial $g(x) \in \mathbb{R}[x]$ is the real symmetric square matrix:

Localizing Matrix: $$\mathbf{M}_{d-d_g}(gy) = \mathrm{E}_\mu[g(x)B_{d-d_g}(x)B^T_{d-d_g}(x)]$$

- B_{d-d_g} : vector of monomial up to order $d - d_g$
- $d_g = \left\lceil \frac{\deg(g(x))}{2} \right\rceil$ the smallest integer larger than $\frac{\deg(g(x))}{2}$

Example: Localizing matrix for moments up to order 4 of a Measure μ in $\mathbb{R}^2$ and polynomial $g(x) = a - x_1^2 - x_2^2$

$d - d_g = 4/2 - 2/2 = 1$

Localizing Matrix

Localizing matrix associated with a sequence of moments up to order $2d$ and a polynomial $g(x) \in \mathbb{R}[x]$ is the real symmetric square matrix:

Localizing Matrix: $$\mathbf{M}_{d-d_g}(gy) = \mathrm{E}_{\mu}[g(x)B_{d-d_g}(x)B^T_{d-d_g}(x)]$$

- B_{d-d_g} : vector of monomial up to order $d - d_g$
- $d_g = \left\lceil \frac{\deg(g(x))}{2} \right\rceil$ the smallest integer larger than $\frac{\deg(g(x))}{2}$

Example: Localizing matrix for moments up to order 4 of a Measure μ in $\mathbb{R}^2$ and polynomial $g(x) = a - x_1^2 - x_2^2$

$d - d_g = 4/2\text{-}2/2=1$

$$\mathbf{M}_1(gy) = \mathrm{E}[g(x)B_1(x)B_1^T(x)] = \mathrm{E}[g(x)\begin{bmatrix}1\\x_1\\x_2\end{bmatrix}\begin{bmatrix}1 & x_1 & x_2\end{bmatrix}] = \mathrm{E}[g(x)\begin{bmatrix}1 & x_1 & x_2\\ x_1 & x_1^2 & x_1x_2\\ x_2 & x_1x_2 & x_2^2\end{bmatrix}] = \mathrm{E}[\begin{bmatrix}g(x) & g(x)x_1 & g(x)x_2\\ g(x)x_1 & g(x)x_1^2 & g(x)x_1x_2\\ g(x)x_2 & g(x)x_1x_2 & g(x)x_2^2\end{bmatrix}]$$

$$= \mathrm{E}[\begin{bmatrix}a - x_1^2 - x_2^2 & ax_1 - x_1^3 - x_1x_2^2 & ax_2 - x_2x_1^2 - x_2^3\\ ax_1 - x_1^3 - x_1x_2^2 & ax_1^2 - x_1^4 - x_1^2x_2^2 & ax_1x_2 - x_1^3x_2 - x_1x_2^3\\ ax_2 - x_2x_1^2 - x_2^3 & ax_1x_2 - x_1^3x_2 - x_1x_2^3 & ax_2^2 - x_2^2x_1^2 - x_2^4\end{bmatrix}] = \begin{bmatrix}ay_{00} - y_{20} - y_{02} & ay_{10} - y_{30} - y_{12} & ay_{01} - y_{21} - y_{03}\\ ay_{10} - y_{30} - y_{12} & ay_{20} - y_{40} - y_{22} & ay_{11} - y_{31} - y_{13}\\ ay_{01} - y_{21} - y_{03} & ay_{11} - y_{31} - y_{13} & ay_{02} - y_{22} - y_{04}\end{bmatrix}$$

Necessary and Sufficient Moment Condition

Let set $\mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ i = 1, \ldots, m\}$ be Archimedean .[1]

Sequence y has a representing measure with support contained in the set $\mathbf{K}$, if and only if, it satisfies:

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d$$

Moment Matrix

$$\mathbf{M}_{d-d_{g_i}}(g_i y) \succcurlyeq 0,\ i = 1, .., m, \quad \forall d$$

Localizing Matrix

- *Chapter 3*: Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

1: e.g., Compact set that includes $M - \|x\|^2 \geq 0, M > 0$

Necessary and Sufficient Moment Condition

Let set $\mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ i = 1, \ldots, m\}$ be Archimedean .[1]

Sequence y has a representing measure with support contained in the set $\mathbf{K}$, if and only if, it satisfies:

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \qquad\qquad \mathbf{M}_{d-d_{g_i}}(g_i y) \succcurlyeq 0,\ i = 1, .., m, \quad \forall d$$

Moment Matrix — Localizing Matrix

- This condition relies on the information of all the moments, i.e., $\mathbf{M}_\infty(y) \succcurlyeq 0,\ \mathbf{M}_\infty(g_i y) \succcurlyeq 0,\ i = 1, .., m.$

- *Chapter 3*: Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

1: e.g., Compact set that includes $M - \|x\|^2 \geq 0, M > 0$

Necessary and Sufficient Moment Condition

Let set $\mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0, \;\; i = 1, ..., m\}$ be Archimedean .[1]

Sequence y has a representing measure with support contained in the set $\mathbf{K}$, if and only if, it satisfies:

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \qquad\qquad \mathbf{M}_{d-d_{g_i}}(g_i y) \succcurlyeq 0, \;\; i = 1, .., m, \quad \forall d$$

Moment Matrix

Localizing Matrix

➢ This condition relies on the information of all the moments, i.e., $\mathbf{M}_\infty(y) \succcurlyeq 0, \;\; \mathbf{M}_\infty(g_i y) \succcurlyeq 0, \;\; i = 1, .., m.$

➢ This is Necessary and Sufficient Condition

$$\boxed{\begin{array}{c} \mu, \;\; supp(\mu) \subset \mathbf{K} \\ y\text{:moments} \end{array}} \iff \boxed{\begin{array}{l} \mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \\ \mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d, \quad i = 1, .., m. \end{array}}$$

- *Chapter 3*: Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

1: e.g., Compact set that includes $M - \|x\|^2 \geq 0, M > 0$

1. Nonlinear Optimization

$$P^* = \underset{x \in \Omega}{\text{minimize}} \quad p(x)$$

- Unconstrained Optimization $\Omega = \mathbb{R}^n$
- Constrained Optimization $\Omega = \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0, \ i = 1, ..., m\}$

2. Optimization in Measure Space

$$P^*_\mu = \underset{\mu \in \mathcal{M}(\Omega)}{\text{minimize}} \ E_\mu[p(x)] = \int p(x) d\mu$$

$$\text{subject to} \int_\Omega d\mu = 1 \longrightarrow \text{Probability Measure}$$

$\mathcal{M}(\Omega)$: space of nonnegative measures on Ω

3. Optimization in Moment Space

$$P^*_{mom} = \underset{y}{\text{minimize}} \ E_\mu[p(x)] = \int p(x) d\mu = \sum_\alpha p_\alpha y_\alpha$$

$$\text{subject to } y_0 = 1 \longrightarrow \text{Probability Measure}$$

y is moment vector of probability distribution $\longleftarrow$ $\text{Moment Matrix } (y) \succcurlyeq 0 \quad \text{Localizing Matrix } (g_i y) \succcurlyeq 0$

1. Nonlinear Optimization

$$P^* = \underset{x \in \Omega}{\text{minimize}} \quad p(x)$$

- Unconstrained Optimization $\Omega = \mathbb{R}^n$
- Constrained Optimization $\Omega = \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ i = 1, ..., m\}$

2. Optimization in Measure Space

$$P^*_\mu = \underset{\mu \in \mathcal{M}(\Omega)}{\text{minimize}} \ \ E_\mu[p(x)] = \int p(x) d\mu$$

$$\text{subject to} \int_\Omega d\mu = 1$$ → Probability Measure

$\mathcal{M}(\Omega)$: space of nonnegative measures on Ω

3. Optimization in Moment Space

$$P^*_{mom} = \underset{y}{\text{minimize}} \ \ E_\mu[p(x)] = \int p(x) d\mu = \sum_\alpha p_\alpha y_\alpha$$

$$\text{subject to } y_0 = 1$$ → Probability Measure

y is moment vector of probability distribution ← $\mathbf{M}_\infty(y) \succcurlyeq 0 \quad \mathbf{M}_\infty(g_i y) \succcurlyeq 0,\ i = 1, .., m$

Dirac Measure

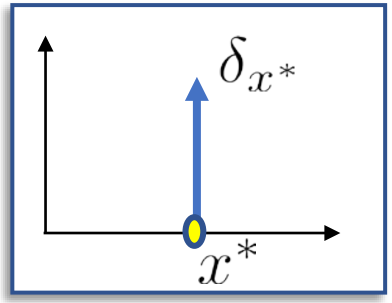
- Optimal solutions in measure space are Dirac measures.

- *Proof of Theorems 5.5, 5.7,1.1:* Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

Dirac Measure

- ➢ Optimal solutions in measure space are Dirac measures.

- ➢ $x^* \in \mathbf{K}\,, p(x^*) = \mathrm{P}^*$:Unique global optimal solution of the original problem.
- ➢ $\mu^* = \delta_{x^*}$:Optimal solution of optimization in measures.



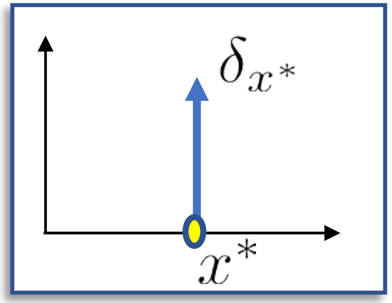
$\mu^* = \delta_{x^*}$

- *Proof of Theorems 5.5, 5.7,1.1:* Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

Dirac Measure

- ➢ Optimal solutions in measure space are Dirac measures.

- ➢ $x^* \in \mathbf{K}$, $p(x^*) = \mathrm{P}^*$:Unique global optimal solution of the original problem.
- ➢ $\mu^* = \delta_{x^*}$:Optimal solution of optimization in measures.



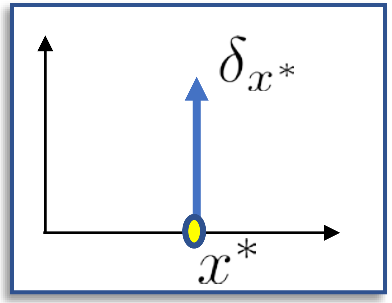
$\mu^* = \delta_{x^*}$

$$\mathrm{P}^*_\mu = \mathrm{E}_\mu[p(x)] = \int p(x)\delta_{x^*}dx = p(x^*) = \mathrm{P}^* \quad \Longrightarrow \quad \mathrm{P}^* = \mathrm{P}^*_\mu$$

- *Proof of Theorems 5.5, 5.7,1.1:* Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

Dirac Measure

- ➢ Optimal solutions in measure space are Dirac measures.

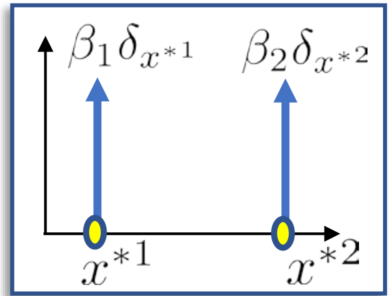


$\mu^* = \delta_{x^*}$

- ➢ $x^* \in \mathbf{K}$, $p(x^*) = \mathrm{P}^*$:Unique global optimal solution of the original problem.
- ➢ $\mu^* = \delta_{x^*}$:Optimal solution of optimization in measures.

$$\mathrm{P}^*_\mu = \mathrm{E}_\mu[p(x)] = \int p(x)\delta_{x^*}dx = p(x^*) = \mathrm{P}^* \quad \Longrightarrow \quad \mathrm{P}^* = \mathrm{P}^*_\mu$$

- ➢ $x^{*i} \in \mathbf{K}, \quad i = 1, ..., r$, $p(x^{*i}) = \mathrm{P}^*$: r global optimal solution of the original problem.
- ➢ $\mu^* = \sum_{i=1}^r \beta_i \delta_{x^{*i}}, \ \ \beta_i > 0, \sum_{i=1}^r \beta_i = 1$:Optimal solution of optimization in measures. (r -atomic measure)



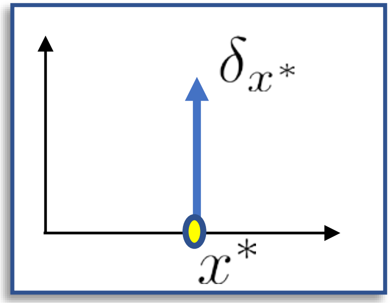
$\mu^* = \beta_1\delta_{x^{*1}} + \beta_2\delta_{x^{*2}}$

- *Proof of Theorems 5.5, 5.7,1.1:* Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

Dirac Measure

- Optimal solutions in measure space are Dirac measures.

- $x^* \in \mathbf{K}\,, p(x^*) = \mathrm{P}^*$:Unique global optimal solution of the original problem.
- $\mu^* = \delta_{x^*}$:Optimal solution of optimization in measures.

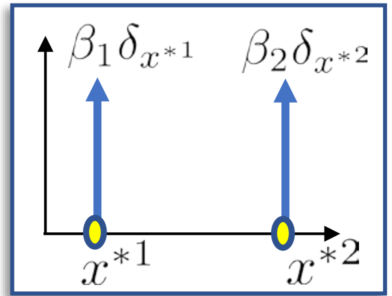


$\mu^* = \delta_{x^*}$

$$\mathrm{P}^*_\mu = \mathrm{E}_\mu[p(x)] = \int p(x)\delta_{x^*}dx = p(x^*) = \mathrm{P}^* \implies \mathrm{P}^* = \mathrm{P}^*_\mu$$

- $x^{*i} \in \mathbf{K}, \quad i = 1, \ldots, r\,, p(x^{*i}) = \mathrm{P}^*$: r global optimal solution of the original problem.
- $\mu^* = \sum_{i=1}^r \beta_i \delta_{x^{*i}}\,, \;\; \beta_i > 0, \sum_{i=1}^r \beta_i = 1$:Optimal solution of optimization in measures. (r -atomic measure)

$$y_0 = 1 \rightarrow \quad y_0 = \sum_i^r \beta_i (x^{*i})^0 \rightarrow \sum_{i=1}^r \beta_i = 1$$



$\mu^* = \beta_1\delta_{x^{*1}} + \beta_2\delta_{x^{*2}}$

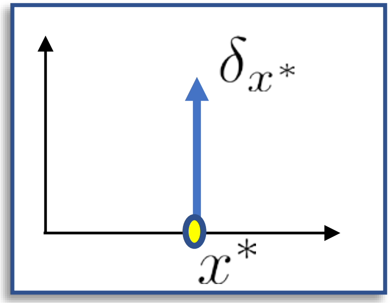
- *Proof of Theorems 5.5, 5.7,1.1:* Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

Dirac Measure

- Optimal solutions in measure space are Dirac measures.

- $x^* \in \mathbf{K}$, $p(x^*) = \mathrm{P}^*$:Unique global optimal solution of the original problem.
- $\mu^* = \delta_{x^*}$:Optimal solution of optimization in measures.

$$\mathrm{P}^*_\mu = \mathrm{E}_\mu[p(x)] = \int p(x)\delta_{x^*}dx = p(x^*) = \mathrm{P}^* \implies \mathrm{P}^* = \mathrm{P}^*_\mu$$

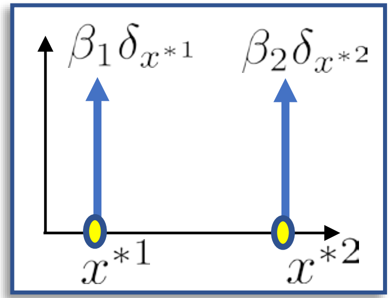


$\mu^* = \delta_{x^*}$

- $x^{*i} \in \mathbf{K},\quad i = 1, \ldots, r$, $p(x^{*i}) = \mathrm{P}^*$: r global optimal solution of the original problem.
- $\mu^* = \sum_{i=1}^r \beta_i \delta_{x^{*i}}$, $\beta_i > 0, \sum_{i=1}^r \beta_i = 1$:Optimal solution of optimization in measures. (r -atomic measure)

$$y_0 = 1 \rightarrow \quad y_0 = \sum_i^r \beta_i (x^{*i})^0 \rightarrow \sum_{i=1}^r \beta_i = 1$$

$$\mathrm{P}^*_\mu = \mathrm{E}_\mu[p(x)] = \int p(x)d\mu^*$$
$$= \int p(x)\left(\sum_{i=1}^r \beta_i \delta_{x^{*i}}\right)dx = \sum_{i=1}^r \beta_i\left(\int p(x)\delta_{x^{*i}}\right)dx = \sum_{i=1}^r \beta_i p(x^{*i}) = \sum_{i=1}^r \beta_i \mathrm{P}^* = \mathrm{P}^*$$



$\mu^* = \beta_1\delta_{x^{*1}} + \beta_2\delta_{x^{*2}}$

- *Proof of Theorems 5.5, 5.7,1.1:* Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

Dirac Measure

- In the moment space, we can identify Dirac measures by looking at the finite sequence of the moments, using the following tests:

- *Theorem 3.7,3.11:* Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

Dirac Measure

- In the moment space, we can identify Dirac measures by looking at the finite sequence of the moments, using the following tests:

Dirac Measure:

Sequence $y = [y_\alpha, \alpha = 0, ..., 2d]$ is the moment sequence of a Dirac measure if and only if

$$M_d(y) \succcurlyeq 0, \quad \mathrm{Rank}(M_d(y)) = 1$$

- *Theorem 3.7,3.11:* Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

Dirac Measure

- In the moment space, we can identify Dirac measures by looking at the finite sequence of the moments, using the following tests:

Dirac Measure:

Sequence $y = [y_\alpha, \alpha = 0, ..., 2d]$ is the moment sequence of a Dirac measure if and only if

$$M_d(y) \succcurlyeq 0, \quad \mathrm{Rank}(M_d(y)) = 1$$

Sequence $y = [y_\alpha, \alpha = 0, ..., 2d]$ is the moment sequence of a Dirac measure with support contained in set $\mathbf{K}$ if and only if

$$M_d(y) \succcurlyeq 0, \;\; M_{d-v}(g_i y) \succcurlyeq 0, i = 1, .., m, \quad \mathrm{Rank}(M_d(y)) = 1$$

$v = \max\{d_{g_i}\}$

- *Theorem 3.7,3.11:* Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

Dirac Measure

- In the moment space, we can identify Dirac measures by looking at the finite sequence of the moments, using the following tests:

Dirac Measure:

Sequence $y = [y_\alpha, \alpha = 0, ..., 2d]$ is the moment sequence of a Dirac measure if and only if

$$M_d(y) \succcurlyeq 0, \quad \mathrm{Rank}(M_d(y)) = 1$$

Sequence $y = [y_\alpha, \alpha = 0, ..., 2d]$ is the moment sequence of a Dirac measure with support contained in set $\mathbf{K}$ if and only if

$$M_d(y) \succcurlyeq 0, \;\; M_{d-v}(g_i y) \succcurlyeq 0, i = 1, .., m, \quad \mathrm{Rank}(M_d(y)) = 1 \qquad v = \max\{d_{g_i}\}$$

Example: $\mu = \delta_2 \Rightarrow y_\alpha = 2^\alpha \quad M_1(y) = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad M_2(y) = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \\ 4 & 8 & 16 \end{bmatrix}$

Moments up to order 2, Rank= 1

Moments up to order 4, Rank= 1

- *Theorem 3.7,3.11:* Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

Dirac Measure

- In the moment space, we can identify Dirac measures by looking at the finite sequence of the moments, using the following tests:

r -atomic measure:

- *Theorem 3.7,3.11:* Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

Dirac Measure

- In the moment space, we can identify Dirac measures by looking at the finite sequence of the moments, using the following tests:

r -atomic measure:

Sequence $y = [y_\alpha, \alpha = 0, ..., 2d]$ is the moment sequence of linear positive combination of r Dirac measure (r -atomic representation) if and only if

$$M_d(y) \succcurlyeq 0, \quad \mathrm{Rank}(M_d(y)) = \mathrm{Rank}(M_{d-1}(y)) = r$$

(M_d is a flat extension of M_{d-1})

- *Theorem 3.7,3.11:* Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

Dirac Measure

- In the moment space, we can identify Dirac measures by looking at the finite sequence of the moments, using the following tests:

r -atomic measure:

Sequence $y = [y_\alpha, \alpha = 0, ..., 2d]$ is the moment sequence of linear positive combination of r Dirac measure (r -atomic representation) if and only if

$$M_d(y) \succcurlyeq 0, \quad \mathrm{Rank}(M_d(y)) = \mathrm{Rank}(M_{d-1}(y)) = r$$

(M_d is a flat extension of M_{d-1})

Sequence $y = [y_\alpha, \alpha = 0, ..., 2d]$ is the moment sequence of linear positive combination of r Dirac measure with support contained in set $\mathbf{K}$,if and only if

$$M_d(y) \succcurlyeq 0, \;\; M_{d-v}(g_i y) \succcurlyeq 0, i = 1, .., m, \quad \mathrm{Rank}(M_d(y)) = \mathrm{Rank}(M_{d-v}(y)) = r \qquad v = \max\{d_{g_i}\}$$

- *Theorem 3.7,3.11:* Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

Dirac Measure

- In the moment space, we can identify Dirac measures by looking at the finite sequence of the moments, using the following tests:

r -atomic measure:

Sequence $y = [y_\alpha, \alpha = 0, ..., 2d]$ is the moment sequence of linear positive combination of r Dirac measure (r -atomic representation) if and only if

$$M_d(y) \succcurlyeq 0, \quad \mathrm{Rank}(M_d(y)) = \mathrm{Rank}(M_{d-1}(y)) = r$$

(M_d is a flat extension of M_{d-1})

Sequence $y = [y_\alpha, \alpha = 0, ..., 2d]$ is the moment sequence of linear positive combination of r Dirac measure with support contained in set $\mathbf{K}$,if and only if

$$M_d(y) \succcurlyeq 0, \; M_{d-v}(g_i y) \succcurlyeq 0, i = 1, .., m, \quad \mathrm{Rank}(M_d(y)) = \mathrm{Rank}(M_{d-v}(y)) = r \qquad v = \max\{d_{g_i}\}$$

Example: $\mu = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$ $\quad y_\alpha = \frac{1}{2}0^\alpha + \frac{1}{2}1^\alpha$ $\quad y = [1, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}]$

$$M_2(y) = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad M_1(y) = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad \mathrm{Rank}(M_2(y)) = \mathrm{Rank}(M_1(y)) = 2$$

- *Theorem 3.7,3.11:* Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

Dirac Measure

- If the solution of the finite SDP in moments satisfies the rank conditions, the obtained moment are moments of Dirac measures.

- Then, we can extract the global optimal solutions of the original problem, as follows:

1: Algorithm 4.2 : Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

Dirac Measure

- If the solution of the finite SDP in moments satisfies the rank conditions, the obtained moment are moments of Dirac measures.

- Then, we can extract the global optimal solutions of the original problem, as follows:

- To obtain the **support point** of a Dirac measure δ_{x^*} (point x^*) from the moment information, we need to look at the moment of order 1.

$$\delta_{x^*} \longrightarrow y_\alpha = {x^*}^\alpha \longrightarrow y_1 = x^*$$

- To obtain the **support points** of a r -atomic measure (linear positive combination of r Dirac measure) from the moment information, we need to solve a linear algebra.
  (relies on moment matrix factorization $\mathrm{M_d(y) = VV^T}$) .[1]

1: Algorithm 4.2 : Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

Topics

1) Overview of Measure and Moment based Optimization

2) Theory of Measure and Moments

3) Measure and Moment based Optimization

4) Examples and GloptiPoly Package

5) Proofs (Moment Conditions)

Measure and Moment based Optimization

1. Nonlinear Optimization

$$P^* = \underset{x \in \Omega}{\text{minimize}} \quad p(x)$$

- Unconstrained Optimization $\Omega = \mathbb{R}^n$
- Constrained Optimization $\Omega = \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0, \ i = 1, ..., m\}$

2. Optimization in Measure Space

$$P^*_\mu = \underset{\mu \in \mathcal{M}(\Omega)}{\text{minimize}} \ \mathrm{E}_\mu[p(x)] = \int p(x) d\mu$$

$$\text{subject to} \int_\Omega d\mu = 1 \longrightarrow \text{Probability Measure}$$

$\mathcal{M}(\Omega)$: space of measures supported on Ω

3. Optimization in Moment Space

$$P^*_{mom} = \underset{y}{\text{minimize}} \ \mathrm{E}_\mu[p(x)] = \int p(x) d\mu = \sum_\alpha p_\alpha y_\alpha$$

$$\text{subject to } y_0 = 1 \longrightarrow \text{Probability Measure}$$

$$\mathbf{M}_\infty(y) \succcurlyeq 0 \quad \mathbf{M}_\infty(g_i y) \succcurlyeq 0, \ i = 1, .., m$$

4. Optimization in Truncated Moment Space

If $y^* = [y^*_\alpha, \alpha = 0, ..., 2d]$ satisfies the rank conditions:
y^* is moment sequence of Dirac measures.
$i) P^{*d}_{mom} = P^*, \ ii) x^*$ can be extracted.
Otherwise $P^{*d}_{mom} \leq P^*$

$$P^{*d}_{mom} = \underset{y_\alpha, \alpha = 0, ..., 2d}{\text{minimize}} \ \mathrm{E}_\mu[p(x)] = \int p(x) d\mu = \sum_\alpha p_\alpha y_\alpha$$

$$\text{subject to } y_0 = 1$$

$$\mathbf{M}_d(y) \succcurlyeq 0 \quad \mathbf{M}_{d-d_{g_i}}(g_i y) \succcurlyeq 0, \ i = 1, .., m$$

4.1 Optimization in Truncated Moment Space $\Omega = \mathbb{R}^n$ (unconstrained Optimization)

- **Decision variable:** $\text{moments: } y_\alpha = \int x^\alpha d\mu$
- **Objective function:** $p(x) = \sum_{\alpha=0}^{2d} p_\alpha x^\alpha \in \mathbb{R}_{2d}[x] \longrightarrow$ Polynomial of order $2d$ (even)
  Otherwise necessarily $p^* = -\infty$

- *Algorithm 5.1, Chapter 7*: Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

4.1 Optimization in Truncated Moment Space $\Omega = \mathbb{R}^n$ (unconstrained Optimization)

- **Decision variable:** $\text{moments: } y_\alpha = \int x^\alpha d\mu$
- **Objective function:** $p(x) = \sum_{\alpha=0}^{2d} p_\alpha x^\alpha \in \mathbb{R}_{2d}[x] \longrightarrow$ Polynomial of order $2d$ (even)
  Otherwise necessarily $p^* = -\infty$

$$
\begin{aligned}
\mathrm{P}^*_{mom} = \underset{y_\alpha, \alpha=0,\dots,2d}{\text{minimize}} \quad & \sum_{\alpha=0}^{2d} p_\alpha y_\alpha \\
\text{subject to} \quad & y_0 = 1 \longrightarrow \int d\mu = 1 \\
& \mathbf{M}_d(y) \succcurlyeq 0
\end{aligned}
$$

- Instead of solving infinite-dimensional SDP, we solve single finite SDP in terms of sequence of the moments up to order $2d$ (degree of $p(x)$).

- *Algorithm 5.1, Chapter 7*: Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

4.1 Optimization in Truncated Moment Space $\Omega = \mathbb{R}^n$ (unconstrained Optimization)

- **Decision variable:** $\text{moments: } y_\alpha = \int x^\alpha d\mu$
- **Objective function:** $p(x) = \sum_{\alpha=0}^{2d} p_\alpha x^\alpha \in \mathbb{R}_{2d}[x] \longrightarrow$ Polynomial of order $2d$ (even)
  Otherwise necessarily $p^* = -\infty$

$$\begin{aligned} \mathrm{P}^*_{mom} = \underset{y_\alpha, \alpha=0,\dots,2d}{\text{minimize}} \quad & \sum_{\alpha=0}^{2d} p_\alpha y_\alpha \\ \text{subject to} \quad & y_0 = 1 \longrightarrow \int d\mu = 1 \\ & \mathbf{M}_d(y) \succcurlyeq 0 \end{aligned}$$

- Instead of solving infinite-dimensional SDP, we solve single finite SDP in terms of sequence of the moments up to order $2d$ (degree of $p(x)$).
- It is similar to SOS programming for unconstrained optimization. Degree of SOS polynomial is determined by the degree of $p(x)$.

$$\mathrm{P} = \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad p(x) \longrightarrow \mathrm{P}_{sos} = \underset{\gamma}{\text{maximize}} \ \gamma \quad \text{subject to, } p(x) - \gamma \in SOS$$

- *Algorithm 5.1, Chapter 7*: Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

4.1 Optimization in Truncated Moment Space $\Omega = \mathbb{R}^n$ (unconstrained Optimization)

- **Decision variable:** $\text{moments: } y_\alpha = \int x^\alpha d\mu$
- **Objective function:** $p(x) = \sum_{\alpha=0}^{2d} p_\alpha x^\alpha \in \mathbb{R}_{2d}[x] \longrightarrow$ Polynomial of order $2d$ (even)
  Otherwise necessarily $p^* = -\infty$

$$
\begin{aligned}
\mathrm{P}^*_{mom} = \underset{y_\alpha, \alpha=0,\ldots,2d}{\text{minimize}} \quad & \sum_{\alpha=0}^{2d} p_\alpha y_\alpha \\
\text{subject to} \quad & y_0 = 1 \longrightarrow \int d\mu = 1 \\
& \mathbf{M}_d(y) \succcurlyeq 0
\end{aligned}
$$

Optimal Solution: Let $y^* = [y^*_\alpha, \alpha = 0, \ldots, 2d]$ be the optimal solution of SDP in moments.

- If $r = \text{Rank } (M_d(y^*)) = \text{Rank } (M_{d-1}(y^*))$, then $\mathrm{P}^{*d}_{mom} = \mathrm{P}^*$ and there are at least r global minimizers which can be extracted by linear algebra (r-atomic measure).
- Otherwise, $\mathrm{P}^*_{mom} \leq \mathrm{P}^*$ (feasible set of finite SDP is bigger that feasible set of infinite SDP)

- *Algorithm 5.1, Chapter 7*: Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

Unconstrained Optimization

$\Omega = \mathbb{R}^n$

$$\mathrm{P}^* = \underset{x\in\Omega}{\text{minimize}} \quad p(x)$$

Infinite LP:

$$\mathrm{P}^*_\mu = \underset{\mu\in\mathcal{M}(\Omega)}{\text{minimize}} \;\; \mathrm{E}_\mu[p(x)] = \int p(x)d\mu$$

$$\text{subject to } \mu(\Omega) = 1$$

$\mathcal{M}(\Omega)$: space of measures supported on Ω

probability measure $:\mu(\Omega) = \int_\Omega d\mu = 1$

SDP

$$\begin{aligned} \mathrm{P}^*_{mom} = \underset{y_\alpha,\alpha=0,\dots,2d}{\text{minimize}} \quad & \sum_{\alpha=0}^{2d} p_\alpha y_\alpha \\ \text{subject to} \quad & y_0 = 1 \\ & \mathbf{M}_d(y) \succcurlyeq 0 \end{aligned}$$

- If $r = \text{Rank } M_d(y^*) = \text{Rank } M_{d-1}(y^*)$

$i)\mathrm{P}^*_{mom} = \mathrm{P}^{*d}$, $ii)x^*$can be extracted by linear algebra

- If $\text{Rank } M_d(y^*) \neq \text{Rank } M_{d-1}(y^*)$

  $\mathrm{P}^*_{mom} \leq \mathrm{P}^*$

4.2 Optimization in Truncated Moment Space $\Omega = \mathbf{K}$ (Constrained Optimization)

- **Decision variable:** moments: $y_\alpha = \int x^\alpha d\mu$
- **Objective function:** $p(x) = \sum_\alpha p_\alpha x^\alpha \in \mathbb{R}[x]$

SDP

$$\begin{aligned} \mathrm{P}^{*d}_{mom} = \underset{y_\alpha, \alpha=0,\dots,2d}{\text{minimize}} \quad & \sum_\alpha p_\alpha y_\alpha \\ \text{subject to} \quad & y_0 = 1 \\ & \mathbf{M}_d(y) \succcurlyeq 0 \\ & \mathbf{M}_{d-d_{g_i}}(g_i y) \succcurlyeq 0, \quad i = 1, \dots, m. \end{aligned}$$

$d_{g_i} = \lceil \frac{deg(g_i(x))}{2} \rceil$

Moments up to order: $2d \geq \max(deg(p(x)), deg(g_i(x)))$

- *Algorithm 5.2, Chapter 7*: Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

4.2 Optimization in Truncated Moment Space $\Omega = \mathbf{K}$ (Constrained Optimization)

- **Decision variable:** moments: $y_\alpha = \int x^\alpha d\mu$
- **Objective function:** $p(x) = \sum_\alpha p_\alpha x^\alpha \in \mathbb{R}[x]$

SDP

$$\begin{aligned} \mathrm{P}^{*d}_{mom} = \underset{y_\alpha, \alpha=0,\ldots,2d}{\text{minimize}} \quad & \sum_\alpha p_\alpha y_\alpha \\ \text{subject to} \quad & y_0 = 1 \\ & \mathbf{M}_d(y) \succcurlyeq 0 \\ & \mathbf{M}_{d-d_{g_i}}(g_i y) \succcurlyeq 0, \quad i = 1, \ldots, m. \end{aligned}$$

$d_{g_i} = \lceil \frac{deg(g_i(x))}{2} \rceil$

Moments up to order: $2d \geq \max(deg(p(x)), deg(g_i(x)))$

Optimal Solution: Let $y^* = [y^*_\alpha, \alpha = 0, \ldots, 2d]$ be the optimal solution of SDP in moments.

- If $r = \text{Rank } M_d(y^*) = \text{Rank } M_{d-v}(y^*)$, then $\mathrm{P}^{*d}_{mom} = \mathrm{P}^*$ and there are at least r global minimizers which can be extracted by linear algebra (r-atomic measure). $v = \max\{d_{g_i}\}$
- If $r = \text{Rank } M_d(y^*) \neq \text{Rank } M_{d-v}(y^*)$, then **increase** relaxation order d by 1 and solve the new obtained SDP.
- Otherwise, $\mathrm{P}^{*d}_{mom} \leq \mathrm{P}^*$

- *Algorithm 5.2, Chapter 7*: Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

4.2 Optimization in Truncated Moment Space $\Omega = \mathbf{K}$ (Constrained Optimization)

- **Decision variable:** moments: $y_\alpha = \int x^\alpha d\mu$
- **Objective function:** $p(x) = \sum_\alpha p_\alpha x^\alpha \in \mathbb{R}[x]$

SDP

$$\begin{aligned} \mathrm{P}^{*d}_{mom} = \underset{y_\alpha, \alpha=0,\ldots,2d}{\text{minimize}} \quad & \sum_\alpha p_\alpha y_\alpha \\ \text{subject to} \quad & y_0 = 1 \\ & \mathbf{M}_d(y) \succcurlyeq 0 \\ & \mathbf{M}_{d-g_i}(g_i y) \succcurlyeq 0, \quad i = 1, \ldots, m. \end{aligned}$$

$d_{g_i} = \lceil \frac{deg(g_i(x))}{2} \rceil$

Moments up to order: $2d \geq \max(deg(p(x)), deg(g_i(x)))$

Optimal Solution: Let $y^* = [y^*_\alpha, \alpha = 0, \ldots, 2d]$ be the optimal solution of SDP in moments.

- If $r = \text{Rank } M_d(y^*) = \text{Rank } M_{d-v}(y^*)$, then $\mathrm{P}^{*d}_{mom} = \mathrm{P}^*$ and there are at least r global minimizers which can be extracted by linear algebra (r-atomic measure). $v = \max\{d_{g_i}\}$
- If $r = \text{Rank } M_d(y^*) \neq \text{Rank } M_{d-v}(y^*)$, then increase relaxation order d by 1 and solve the new obtained SDP.
- Otherwise, $\mathrm{P}^{*d}_{mom} \leq \mathrm{P}^*$

Theorem: as $d \to \infty$ (number of moments in SDP increases), $\mathrm{P}^{*d}_{mom} \uparrow \mathrm{P}^*$

- *Theorem 5.6, Chapter 7*: Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

Finite Convergence

$$\exists \; d^* \quad \mathrm{P}^{*d^*}_{mom} = \mathrm{P}^* \quad d \geq d^*$$

- Typically, finite convergence takes place for some (usually small) d.
- Sufficient condition for finite convergence, is the rank condition of the moments.

- Nie, J. Optimality Conditions and Finite Convergence of Lasserre's Hierarchy, Math. Program. Ser. A 146, 97–121, 2014.
- Jean Lasserre "The Moment-SOS hierarchy", 2018

Finite Convergence

$$\exists\ d^* \quad \mathrm{P}^{*d^*}_{mom} = \mathrm{P}^* \quad d \geq d^*$$

- Typically, finite convergence takes place for some (usually small) d.
- Sufficient condition for finite convergence, is the rank condition of the moments.
- For the sub-class of *Convex Optimization* Problems where $p, -g_i \;\; i = 1, \ldots, m$ are *SOS-Convex* polynomials, finite convergence takes place at the first step of the hierarchy.
  $\text{Relaxation Order: } 2d = \max(deg(p(x)), deg(g_i(x)))$

- Nie, J. Optimality Conditions and Finite Convergence of Lasserre's Hierarchy, Math. Program. Ser. A 146, 97–121, 2014.
- Jean Lasserre "The Moment-SOS hierarchy", 2018

4.2 Optimization in Truncated Moment Space $\Omega = \mathbf{K}$ (Constrained Optimization)

Finite Convergence

$$\exists \; d^* \quad \mathrm{P}^{*d^*}_{mom} = \mathrm{P}^* \quad d \geq d^*$$

- Typically, finite convergence takes place for some (usually small) d.
- Sufficient condition for finite convergence, is the rank condition of the moments.
- For the sub-class of *Convex Optimization* Problems where $p, -g_i \;\; i = 1, ..., m$ are *SOS-Convex* polynomials, finite convergence takes place at the first step of the hierarchy.

  $\text{Relaxation Order: } 2d = \max(deg(p(x)), deg(g_i(x)))$

  $\nabla^2 p \succcurlyeq 0$ *(Convexity condition)*

  $\nabla^2 p = L(x) L^T(x)$ *(SOS-Convexity condition)*

- Nie, J. Optimality Conditions and Finite Convergence of Lasserre's Hierarchy, Math. Program. Ser. A 146, 97–121, 2014.
- Jean Lasserre "The Moment-SOS hierarchy", 2018

4.2 Optimization in Truncated Moment Space $\Omega = \mathbf{K}$ (Constrained Optimization)

Finite Convergence

$$\exists\ d^* \quad \mathrm{P}^{*d^*}_{mom} = \mathrm{P}^* \quad d \geq d^*$$

- Typically, finite convergence takes place for some (usually small) d.
- Sufficient condition for finite convergence, is the rank condition of the moments.
- For the sub-class of *Convex Optimization* Problems where $p, -g_i \ \ i = 1, ..., m$ are *SOS-Convex* polynomials, finite convergence takes place at the first step of the hierarchy.

  $\text{Relaxation Order: } 2d = \max(deg(p(x)), deg(g_i(x)))$

$$\mathrm{P}^* = \underset{x \in \Omega}{\text{minimize}} \quad p(x)$$
$$\Omega = \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ \ i = 1, ..., m\}$$

$\Rightarrow$ $p, -g_i \ \ i = 1, ..., m$ *SOS-Convex* $\Rightarrow$ **This implies the Convexity of the original optimization.**

- Nie, J. Optimality Conditions and Finite Convergence of Lasserre's Hierarchy, Math. Program. Ser. A 146, 97–121, 2014.
- Jean Lasserre "The Moment-SOS hierarchy", 2018

Unconstrained Optimization

$\Omega = \mathbb{R}^n$

$$\mathrm{P}^* = \underset{x \in \Omega}{\text{minimize}} \quad p(x)$$

Constrained Optimization

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

Infinite LP:

$$\mathrm{P}^*_\mu = \underset{\mu \in \mathcal{M}(\Omega)}{\text{minimize}} \ \ \mathrm{E}_\mu[p(x)] = \int p(x) d\mu$$

$$\text{subject to } \mu(\Omega) = 1$$

$\mathcal{M}(\Omega)$: space of nonnegative measures on Ω

probability measure $: \mu(\Omega) = \int_\Omega d\mu = 1$

SDP

$$\mathrm{P}^*_{mom} = \underset{y_\alpha, \alpha = 0, \ldots, 2d}{\text{minimize}} \quad \sum_{\alpha=0}^{2d} p_\alpha y_\alpha$$

$$\text{subject to} \quad y_0 = 1$$

$$\mathbf{M}_d(y) \succcurlyeq 0$$

$2d = deg(p(x))$

- If $r = \text{Rank } M_d(y^*) = \text{Rank } M_{d-1}(y^*)$
  $i)\mathrm{P}^*_{mom} = \mathrm{P}^*, \ \ ii) x^*$ can be extracted by linear algebra
- If $\text{Rank } M_d(y^*) \neq \text{Rank } M_{d-1}(y^*)$
  $\mathrm{P}^*_{mom} \leq \mathrm{P}^*$

SDP

$$\mathrm{P}^{*d}_{mom} = \underset{y_\alpha, \alpha = 0, \ldots, 2d}{\text{minimize}} \quad \sum_\alpha p_\alpha y_\alpha$$

$$\text{subject to} \quad y_0 = 1$$

$$\mathbf{M}_d(y) \succcurlyeq 0$$

$$\mathbf{M}_{d-d_{g_i}}(g_i y) \succcurlyeq 0, \quad i = 1, \ldots, m.$$

$d_{g_i} = \lceil \frac{deg(g_i(x))}{2} \rceil$

$2d \geq \max\left(deg(p(x)), deg(g_i(x))\right)$

- If $r = \text{Rank } M_d(y^*) = \text{Rank } M_{d-v}(y^*)$ $\quad v = \max\{d_{g_i}\}$
  $i)\mathrm{P}^*_{mom} = \mathrm{P}^*, \ \ ii) x^*$ can be extracted by linear algebra
- If $\text{Rank } M_d(y^*) \neq \text{Rank } M_{d-v}(y^*)$, increas d
- Otherwise $\mathrm{P}^{*d}_{mom} \leq \mathrm{P}^*$

Topics

1) Overview of Measure and Moment based Optimization

2) Theory of Measure and Moments

3) Measure and Moment based Optimization

4) Examples and GloptiPoly Package

5) Proofs (Moment Conditions)

Gloptipoly: MATLAB based packaged for moment-measure based optimization.

http://homepages.laas.fr/henrion/software/gloptipoly3/

Input: Nonlinear Optimization

- Generates moment based Semidefinite Program (SDP)
- Solves the SDP using SDP solvers
- Checks the rank condition and extracts the solution

SDP solvers: e.g.,

MOSEK https://www.mosek.com

SEDUMI http://sedumi.ie.lehigh.edu

SDPT3 http://www.math.nus.edu.sg/~mattohkc/sdpt3.html

Rely on **interior point** methods

Example 1: Unconstrained Optimization

$$P^* = \underset{x\in\mathbb{R}^2}{\text{minimize}} \quad 3 + 2x_1 + 2x_2 + 3x_1^2 + 2x_1x_2 + 3x_2^2 + x_1^4 + x_2^4$$

$deg(p(x)) = 4 \Longrightarrow$ SDP with moments up to order 4 (d=2).

Variables of SDP: $y = [y_{00}|y_{10}, y_{01}|y_{20}, y_{11}, y_{02}|y_{30}, y_{21}, y_{12}, y_{03}|y_{40}, y_{31}, y_{22}, y_{13}, y_{04}]$

$$P^{*2}_{mom} = \underset{y}{\text{minimize}} \quad 3 + 2y_{10} + 2y_{01} + 3y_{20} + 2y_{11} + 3y_{02} + y_{40} + y_{04}$$

$$\text{subject to}$$

$$y_{00} = 1 \qquad \mathbf{M}_2(y) = \begin{bmatrix} y_{00} & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ y_{01} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{bmatrix} \succcurlyeq 0$$

Status=1 (Rank condition holds) : $P^{*2}_{mom} = P^* = 2.5074 \qquad x_1^* = -0.2428,\ x_2^* = -0.2428$

- Objective function admits SOS representation $(x_1^2+1)^2 + (x_2^2+1)^2 + (x_1+x_2+1)^2$

Moment Programming in GloptiPoly

```
mset('yalmip',true);mset(sdpsettings('solver','mosek'));
mpol x1 x2
p =3+2*x1+2*x2+3*x1^2+2*x1*x2+3*x2^2+x1^4+x2^4;
P = msdp(min(p))
[status,obj] = msol(P)
```

→ SDP solver
→ variables x_1, x_2
→ $p(x)$
→ genrate moment SDP
→ solve moment SDP

Output:

status==1:
The moment SDP is solved successfully, and Rank conditions are satisfied. Hence, GloptiPoly can extract the global optimal solutions.
obj = the global minimum , $x^* =$ double(x)

status==0:
The moment SDP is solved successfully But Rank conditions are Not satisfied. Hence, GloptiPoly can NOT extract the global optimal solutions.
obj is lover bound of the global minimum.

status==-1: The moment SDP could NOT be solved (unbounded SDP).

https://github.com/jasour/rarnop19/blob/master/Lecture4_Moment_NonlinearOptimization/Example_1.m

Moment Programming in GloptiPoly

```
mset('yalmip',true);mset(sdpsettings('solver','mosek'));
mpol x1 x2
p =3+2*x1+2*x2+3*x1^2+2*x1*x2+3*x2^2+x1^4+x2^4;
P = msdp(min(p))
[status,obj] = msol(P)
```

→ SDP solver
→ variables x_1, x_2
→ $p(x)$
→ genrate moment SDP
→ solve moment SDP

GloptiPoly Output: Description of the Moment SDP

Measure 1
  Degree = 4 → $\deg\big(p(x)\big) = 2d$
  Variables = 2 → x_1, x_2
  Moments = 15 → $S_{n,2d} = \binom{n+2d}{n} = \frac{(n+2d)!}{n!2d!}$
Relaxation order = 2 → d
Mass of measure 1 set to one → $y_0 = 1$
Total number of moments = 15 → $S_{n,2d} = \binom{n+2d}{n} = \frac{(n+2d)!}{n!2d!}$
Perform moment substitutions
Number of moments after substitutions = 14 → $y_0 = 1$

https://github.com/jasour/rarnop19/blob/master/Lecture4_Moment_NonlinearOptimization/Example_1.m

Moment Programming in GloptiPoly

```
mset('yalmip',true);mset(sdpsettings('solver','mosek'));
mpol x1 x2
p =3+2*x1+2*x2+3*x1^2+2*x1*x2+3*x2^2+x1^4+x2^4;
P = msdp(min(p))
[status,obj] = msol(P)
```

→ SDP solver
→ variables x_1, x_2
→ $p(x)$
→ genrate moment SDP
→ solve moment SDP

Obtained Results

```
double([x1 x2])
y=double(mvec(meas));
M=double(mmat(meas));
```

→ global optimal solution (If status=1)
→ obtained moment vector (if status $\neq -1$)
→ moment matrix of the moments (if status $\neq -1$)

Rank Test (if status $\neq -1$)

```
d2= deg(p)/2;
B_d2=mmon([x1 x2],d2);
M_d2=double(mom(B_d2*B_d2'));
R_d2=eig(M_d2);
```

Moment Matrix of order d : $M_d = \mathrm{E}[B_d B_d^T]$
Rank: nonzero eigenvalues

$\Rightarrow$ $r = \text{Rank } M_d(y^*) = \text{Rank } M_{d-1}(y^*)$

```
d1= d2-1;
B_d1=mmon([x1 x2],d1);
M_d1=double(mom(B_d1*B_d1'));
R_d1=eig(M_d1);
```

Moment Matrix of order $d-1$: $M_{d-1} = \mathrm{E}[B_{d-1} B_{d-1}^T]$
Rank: nonzero eigenvalues

https://github.com/jasour/rarnop19/blob/master/Lecture4_Moment_NonlinearOptimization/Example_1.m

Example 2: Unconstrained Optimization

$$P^* = \underset{x\in\mathbb{R}^2}{\text{minimize}} \quad -x_1^2x_2^2 + x_1^4x_2^2 + x_1^2x_2^4$$

$$P^* = -1/27,\ x^* = [\pm\sqrt{3}/3, \pm\sqrt{3}/3]$$

$deg(p(x)) = 6$ ⇒ SDP with moments up to order 6 (d=3)

Variables of SDP: $y = [y_{00}|y_{10}, y_{01}|y_{20}, y_{11}, y_{02}|y_{30}, y_{21}, y_{12}, y_{03}|y_{40}, y_{31}, y_{22}, y_{13}, y_{04}|$
$y_{50}|y_{41}, y_{32}, y_{23}, y_{14}, y_{05}|y_{60}, y_{51}, y_{42}, y_{33}|y_{24}, y_{15}, y_{06}]$

$$P^{*2}_{mom} = \underset{y}{\text{minimize}} \quad -y_{22} + y_{42} + y_{24}$$

subject to

$$y_{00} = 1 \qquad \mathbf{M}_2(y) = \begin{bmatrix} y_{00} & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ y_{01} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{bmatrix} \succcurlyeq 0$$

Status=-1 : moment SDP could NOT be solved (unbounded SDP).

https://github.com/jasour/rarnop19/blob/master/Lecture4_Moment_NonlinearOptimization/Example_2.m

- Objective function is nonnegative but does not admit SOS representation

Example 3: Constrained Optimization

$$\mathrm{P}^* = \underset{x \in \mathbb{R}^2}{\text{minimize}} \quad -x_1$$
$$\text{subject to} \quad x \in \mathbf{K} = \{x \in \mathbb{R}^2 : 3 - 2x_2 - x_1^2 - x_2^2 \geq 0,\ -x_1 - x_2 - x_1 x_2 \geq 0,\ 1 + x_1 x_2 \geq 0\}$$

$$\max(deg(p(x)), deg(g_i(x))) = 2$$

Minimum relaxation order of SDP: 2 (d=1)
SDP with moments up to order 2.

$$\mathrm{P}^{*1}_{mom} = \underset{y}{\text{minimize}} \quad -y_{10}$$

subject to

$$\mathbf{M}_1(y) = \begin{bmatrix} y_{00} & y_{10} & y_{01} \\ y_{10} & y_{20} & y_{11} \\ y_{01} & y_{11} & y_{02} \end{bmatrix} \succcurlyeq 0$$

$$y_{00} = 1$$
$$\mathbf{M}_{d-d_i}(g_1 y) = 3 - 2y_{01} - y_{20} - y_{02} \geq 0,$$
$$\mathbf{M}_{d-d_i}(g_2 y) = -y_{10} - y_{01} - y_{11} \geq 0,$$
$$\mathbf{M}_{d-d_i}(g_3 y) = 1 + y_{11} \geq 0$$

Status=0 : The moment SDP is solved successfully But Rank conditions are Not satisfied. Hence, GloptiPoly can NOT extract the global optimal solutions.

$$\mathrm{P}^{*1}_{mom} = -1.7808 \leq \mathrm{P}^*$$

https://github.com/jasour/rarnop19/blob/master/Lecture4_Moment_NonlinearOptimization/Example_3.m

$$P^* = \underset{x\in\mathbb{R}^2}{\text{minimize}} \quad -x_1$$

$$\text{subject to} \quad x \in \mathbf{K} = \{x \in \mathbb{R}^2 : 3 - 2x_2 - x_1^2 - x_2^2 \geq 0,\ -x_1 - x_2 - x_1x_2 \geq 0,\ 1 + x_1x_2 \geq 0\}$$

Moment Programming in GloptiPoly

```
mset('yalmip',true);mset(sdpsettings('solver','mosek'));
mpol x1 x2
p = (1+x1*x2)^2-x1*x2+(1-x2)^2;
g=[3-2*x2-x1^2-x2^2;-x1-x2-x1*x2;1+x1*x2];
P = msdp(min(p), g>=0, d)

[status,obj] = msol(P)
```

- `mset('yalmip',true);mset(sdpsettings('solver','mosek'));` → SDP solver
- `mpol x1 x2` → variables x_1, x_2
- `p = (1+x1*x2)^2-x1*x2+(1-x2)^2;` → $p(x)$
- `g=[3-2*x2-x1^2-x2^2;-x1-x2-x1*x2;1+x1*x2];` → $\mathbf{K}$
- `P = msdp(min(p), g>=0, d)` → genrate moment SDP relaxation of d, $2d \geq \max(deg(p(x)), deg(g_i(x)))$
- `[status,obj] = msol(P)` → solve moment SDP

https://github.com/jasour/rarnop19/blob/master/Lecture4_Moment_NonlinearOptimization/Example_3.m

Moment Programming in GloptiPoly

```
mset('yalmip',true);mset(sdpsettings('solver','mosek'));
mpol x1 x2
p = (1+x1*x2)^2-x1*x2+(1-x2)^2;
g=[3-2*x2-x1^2-x2^2;-x1-x2-x1*x2;1+x1*x2];
P = msdp(min(p), g>=0, d)

[status,obj] = msol(P)
```

- `mset('yalmip',true);mset(sdpsettings('solver','mosek'));` → SDP solver
- `mpol x1 x2` → variables x_1, x_2
- `p = (1+x1*x2)^2-x1*x2+(1-x2)^2;` → $p(x)$
- `g=[3-2*x2-x1^2-x2^2;-x1-x2-x1*x2;1+x1*x2];` → $\mathbf{K}$
- `P = msdp(min(p), g>=0, d)` → genrate moment SDP relaxation of d, $2d \geq \max(deg(p(x)), deg(g_i(x)))$
- `[status,obj] = msol(P)` → solve moment SDP

GloptiPoly Output: Description of the Moment SDP

Number of support constraints = 3 → $g_i(x), i = 1,2,3$
Measure 1
Degree = 2 → $d = \max(\ \deg(p(x)), \deg(g_i(x))\ )$
Variables = 2 → x_1, x_2
Moments = 6 → $S_{n,2d} = \binom{n+2d}{n} = \frac{(n+2d)!}{n!2d!}$
Relaxation order = 1 → $\lceil d/2 \rceil$
Mass of measure 1 set to one → $y_0 = 1$
Total number of moments = 6 → $S_{n,2d} = \binom{n+2d}{n} = \frac{(n+2d)!}{n!2d!}$
Perform moment substitutions
Number of moments after substitutions = 5 → $y_0 = 1$

https://github.com/jasour/rarnop19/blob/master/Lecture4_Moment_NonlinearOptimization/Example_3.m

Moment Programming in GloptiPoly

```
mset('yalmip',true);mset(sdpsettings('solver','mosek'));
mpol x1 x2
p = (1+x1*x2)^2-x1*x2+(1-x2)^2;
g=[3-2*x2-x1^2-x2^2;-x1-x2-x1*x2;1+x1*x2];
P = msdp(min(p), g>=0, d)

[status,obj] = msol(P)
```

→ SDP solver

→ variables x_1, x_2

→ $p(x)$

→ $\mathbf{K}$

→ genrate moment SDP relaxation of d
$2d \geq \max(deg(p(x)), deg(g_i(x)))$

→ solve moment SDP

Obtained Results

```
double([x1 x2])
y=double(mvec(meas));
M=double(mmat(meas));
```

→ global optimal solution (If status=1)

→ obtained moment vector (if status $\neq -1$)

→ moment matrix of the moments (if status $\neq -1$)

Rank Test (if status $\neq -1$)

```
d2= d;
B_d2=mmon([x1 x2],d2);
M_d2=double(mom(B_d2*B_d2'));
R_d2=eig(M_d2);
```

Moment Matrix of order $d : M_d = \mathrm{E}[B_d B_d^T]$
Rank: nonzero eigenvalues

```
d1= d2-ceil(deg(g)/2);
B_d1=mmon([x1 x2],d1);
M_d1=double(mom(B_d1*B_d1'));
R_d1=eig(M_d1);
```

Moment Matrix of order $d-v : M_{d-v} = \mathrm{E}[B_{d-v} B_{d-v}^T]$
Rank: nonzero eigenvalues

⇒ $r = \text{Rank } M_d(y^*) = \text{Rank } M_{d-v}(y^*)$

$v = \max\{d_{g_i}\} \quad d_{g_i} = \lceil \frac{deg(g_i(x))}{2} \rceil$

https://github.com/jasour/rarnop19/blob/master/Lecture4_Moment_NonlinearOptimization/Example_3.m

Example 4: Constrained Optimization

We solve the Optimization of Example 3 with relaxation order $d = 2$

SDP with moments up to order 2d

$$\mathrm{P}^{*1}_{mom} = \underset{y}{\text{minimize}} \quad -y_{10}$$

$$\text{subject to} \quad y_{00} = 1$$

$$\mathbf{M}_2(y) = \begin{bmatrix} y_{00} & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ y_{01} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{bmatrix} \succcurlyeq 0 \qquad \mathbf{M}_2(g_1 y) = \begin{bmatrix} 3y_{00} - 2y_{01} - y_{20} - y_{02} & 3y_{10} - 2y_{11} - y_{30} - y_{12} & 3y_{01} - 2y_{02} - y_{21} - y_{03} \\ 3y_{10} - 2y_{11} - y_{30} - y_{12} & 3y_{20} - 2y_{21} - y_{40} - y_{22} & 3y_{11} - 2y_{12} - y_{31} - y_{13} \\ 3y_{01} - 2y_{02} - y_{21} - y_{03} & 3y_{11} - 2y_{12} - y_{31} - y_{13} & 3y_{02} - 2y_{03} - y_{22} - y_{04} \end{bmatrix} \succcurlyeq 0$$

$$\mathbf{M}_2(g_2 y) = \begin{bmatrix} -y_{10} - y_{01} - y_{11} & -y_{20} - y_{11} - y_{21} & -y_{11} - y_{02} - y_{12} \\ -y_{20} - y_{11} - y_{21} & -y_{30} - y_{21} - y_{31} & -y_{21} - y_{12} - y_{22} \\ -y_{11} - y_{02} - y_{12} & -y_{21} - y_{12} - y_{22} & -y_{12} - y_{03} - y_{13} \end{bmatrix} \succcurlyeq 0 \qquad \mathbf{M}_2(g_3 y) = \begin{bmatrix} y_{00} + y_{11} & y_{10} + y_{21} & y_{01} + y_{12} \\ y_{10} + y_{21} & y_{20} + y_{31} & y_{11} + y_{22} \\ y_{01} + y_{12} & y_{11} + y_{22} & y_{02} + y_{13} \end{bmatrix} \succcurlyeq 0$$

Status=1 (rank condition holds) $\quad \mathrm{P}^{*1}_{mom} = -1.6180 = \mathrm{P}^* \qquad x_1^* = 1.6810, \; x_2^* = -0.6180$

https://github.com/jasour/rarnop19/blob/master/Lecture4_Moment_NonlinearOptimization/Example_4.m

Example 5: Constrained Optimization

$$\mathrm{P}^* = \underset{x\in\mathbb{R}^3}{\text{minimize}} \quad -2x_1 + x_2 - x_3$$

$$\text{subject to}$$

$$24 - 20x_1 + 9x_2 - 13x_3 + 4x_1^2 - 4x_1x_2 + 4x_1x_3 + 2x_2^2 - 2x_2x_3 + 2x_3^2 \geq 0,$$

$$x_1 + x_2 + x_3 \leq 4,$$

$$3x_2 + x_3 \leq 6$$

$$0 \leq x_1 \leq 2, \quad 0 \leq x_2, \quad 0 \leq x_3 \leq 3$$

$$\max(deg(p(x)), deg(g_i(x)) = 2 \longrightarrow 2d \geq 2$$

GloptiPoly with SeDumi SDP solver:

$d = 1$, Status=0 $\quad \mathrm{P}^{*1}_{mom} = -6 \leq \mathrm{P}^*$

$d = 2$, Status=0 $\quad \mathrm{P}^{*2}_{mom} = -5.6923 \leq \mathrm{P}^*$

$d = 3$, Status=0 $\quad \mathrm{P}^{*3}_{mom} = -4.0683 \leq \mathrm{P}^*$

$d = 4$, Status=1 $\quad \mathrm{P}^{*4}_{mom} = -4 = \mathrm{P}^* \quad x_1^* = 0.5,\ x_2^* = 0,\ x_3^* = 0 \quad x_1^* = 2,\ x_2^* = 0,\ x_3^* = 0$

Note that for $d \geq 4$ $\quad \mathrm{P}^{*d}_{mom} = -4 = \mathrm{P}^*$

https://github.com/jasour/rarnop19/blob/master/Lecture4_Moment_NonlinearOptimization/Example_5.m

Topics

1) Overview of Measure and Moment based Optimization

2) Theory of Measure and Moments

3) Measure and Moment based Optimization

4) Examples and GloptiPoly Package

5) Proofs (Moment Conditions)

1

Optimization problems in x and in μ are equivalent, $\mathrm{P}^* = \mathrm{P}^*_\mu$

2

Moments of every measure μ in $\mathbb{R}^n$ satisfies the moment condition:

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d$$

3

Moments of every measure μ on the set $\mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0 \; |_{i=1}^m\}$ satisfies the moment condition:

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \qquad \mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d, \quad i = 1, .., m.$$

4

Any infinite sequence y that satisfies $\mathbf{M}_d(y) \succcurlyeq 0, \mathbf{M}_d(g_i y) \succcurlyeq 0, \; \forall d, \; i = 1, .., m$ are the moment seq. of measure μ supported on compact (Archimedean) set $\mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0 \; |_{i=1}^m\}$

1

Optimization problems in x and in μ are equivalent, $\mathrm{P}^* = \mathrm{P}^*_\mu$

- *Chapter 7*: Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

1

Optimization problems in x and in μ are equivalent, $\mathrm{P}^* = \mathrm{P}^*_\mu$

To show $\mathrm{P}^* = \mathrm{P}^*_\mu$

1 We first show that $\mathrm{P}^* \leq \mathrm{P}^*_\mu$

2 We show that $\mathrm{P}^* \geq \mathrm{P}^*_\mu$

1 2 $\Longrightarrow$ $\mathrm{P}^* = \mathrm{P}^*_\mu$

- *Chapter 7*: Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

1

Optimization problems in x and in μ are equivalent, $\mathrm{P}^* = \mathrm{P}^*_\mu$

To show $\mathrm{P}^* = \mathrm{P}^*_\mu$

1 We first show that $\mathrm{P}^* \leq \mathrm{P}^*_\mu$

$$\mathrm{P}^* \leq p(x) \quad \forall x \in \Omega$$

2 We show that $\mathrm{P}^* \geq \mathrm{P}^*_\mu$

1 2 $\Longrightarrow$ $\mathrm{P}^* = \mathrm{P}^*_\mu$

- *Chapter 7*: Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

1

Optimization problems in x and in μ are equivalent, $\mathrm{P}^* = \mathrm{P}^*_\mu$

To show $\mathrm{P}^* = \mathrm{P}^*_\mu$

1 We first show that $\mathrm{P}^* \leq \mathrm{P}^*_\mu$

$$\mathrm{P}^* \leq p(x) \quad \forall x \in \Omega \longrightarrow \mathrm{E}_\mu[\mathrm{P}^*] \leq \mathrm{E}_\mu[p(x)] \Longrightarrow \mathrm{P}^* \leq \mathrm{P}^*_\mu$$

2 We show that $\mathrm{P}^* \geq \mathrm{P}^*_\mu$

1 2 $\Longrightarrow \mathrm{P}^* = \mathrm{P}^*_\mu$

- *Chapter 7*: Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

1

Optimization problems in x and in μ are equivalent, $\mathrm{P}^* = \mathrm{P}^*_\mu$

To show $\mathrm{P}^* = \mathrm{P}^*_\mu$

1 We first show that $\mathrm{P}^* \leq \mathrm{P}^*_\mu$

$$\mathrm{P}^* \leq p(x) \quad \forall x \in \Omega \longrightarrow \mathrm{E}_\mu[\mathrm{P}^*] \leq \mathrm{E}_\mu[p(x)] \implies \mathrm{P}^* \leq \mathrm{P}^*_\mu$$

2 We show that $\mathrm{P}^* \geq \mathrm{P}^*_\mu$

- For every feasible solution $x \in \Omega$, we associate the Dirac measure δ_x.
- $\delta_x \in \mathcal{M}(\Omega)$ is a feasible solution of optimization problem in μ

1 2 $\implies \mathrm{P}^* = \mathrm{P}^*_\mu$

- *Chapter 7*: Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

1

Optimization problems in x and in μ are equivalent, $\mathrm{P}^* = \mathrm{P}^*_\mu$

To show $\mathrm{P}^* = \mathrm{P}^*_\mu$

1 We first show that $\mathrm{P}^* \leq \mathrm{P}^*_\mu$

$$\mathrm{P}^* \leq p(x) \quad \forall x \in \Omega \quad \longrightarrow \quad \mathrm{E}_\mu[\mathrm{P}^*] \leq \mathrm{E}_\mu[p(x)] \quad \Longrightarrow \quad \mathrm{P}^* \leq \mathrm{P}^*_\mu$$

2 We show that $\mathrm{P}^* \geq \mathrm{P}^*_\mu$

- For every feasible solution $x \in \Omega$, we associate the Dirac measure δ_x.
- $\delta_x \in \mathcal{M}(\Omega)$ is a feasible solution of optimization problem in μ

$$\mathrm{P}^* = \underset{x\in\Omega}{\text{minimize}} \ \ p(x) = \underset{\delta_x\in\mathcal{M}(\Omega)}{\text{minimize}} \ \ \mathrm{E}[p(x)]$$

1 2 $\Longrightarrow$ $\mathrm{P}^* = \mathrm{P}^*_\mu$

- *Chapter 7*: Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

1

Optimization problems in x and in μ are equivalent, $\mathrm{P}^* = \mathrm{P}_\mu^*$

To show $\mathrm{P}^* = \mathrm{P}_\mu^*$

1 We first show that $\mathrm{P}^* \leq \mathrm{P}_\mu^*$

$$\mathrm{P}^* \leq p(x) \quad \forall x \in \Omega \longrightarrow \mathrm{E}_\mu[\mathrm{P}^*] \leq \mathrm{E}_\mu[p(x)] \Longrightarrow \mathrm{P}^* \leq \mathrm{P}_\mu^*$$

2 We show that $\mathrm{P}^* \geq \mathrm{P}_\mu^*$

- For every feasible solution $x \in \Omega$, we associate the Dirac measure δ_x.
- $\delta_x \in \mathcal{M}(\Omega)$ is a feasible solution of optimization problem in μ

$$\mathrm{P}^* = \underset{x \in \Omega}{\text{minimize}} \ p(x) = \underset{\delta_x \in \mathcal{M}(\Omega)}{\text{minimize}} \ \mathrm{E}[p(x)] \geq \mathrm{P}_\mu^* = \underbrace{\underset{\mu \in \mathcal{M}(\Omega)}{\text{minimize}} \ \mathrm{E}[p(x)]}_{\text{Bigger feasible Set}} \Longrightarrow \mathrm{P}^* \geq \mathrm{P}_\mu^*$$

1 2 $\Longrightarrow \mathrm{P}^* = \mathrm{P}_\mu^*$

- *Chapter 7*: Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

2

Moments of every measure μ in $\mathbb{R}^n$ satisfies the moment condition:

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d$$

2

Moments of every measure μ in $\mathbb{R}^n$ satisfies the moment condition:

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d$$

We want to show that given a measure and its moments sequence, μ, y $\Longrightarrow$ $\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d$

- *Eq 3.11*: Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

2

Moments of every measure μ in $\mathbb{R}^n$ satisfies the moment condition:

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d$$

We want to show that given a measure and its moments sequence, μ, y $\Longrightarrow$ $\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d$

- We need to show that $\mathbf{p}^T \mathbf{M}_d(y) \mathbf{p} \geq 0, \quad \forall$ vector: $\mathbf{p} \in \mathbb{R}^{S_{n,d}}$ (definition of PSD matrix)

2

Moments of every measure μ in $\mathbb{R}^n$ satisfies the moment condition:

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d$$

We want to show that given a measure and its moments sequence, $\mu, y \Longrightarrow \mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d$

- We need to show that $\mathbf{p}^T \mathbf{M}_d(y) \mathbf{p} \geq 0, \quad \forall \text{ vector: } \mathbf{p} \in \mathbb{R}^{S_{n,d}}$ (definition of PSD matrix)

$$\mathbf{p}^T \mathbf{M}_d(y) \mathbf{p} = \mathbf{p}^T \underbrace{\mathrm{E}[B_d(x) B_d^T(x)]}_{} \mathbf{p} = \mathrm{E}[\mathbf{p}^T B_d(x) B_d^T(x) \mathbf{p}]$$

(definition of Moment matrix)

2

Moments of every measure μ in $\mathbb{R}^n$ satisfies the moment condition:

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d$$

We want to show that given a measure and its moments sequence, $\mu, y \Longrightarrow \mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d$

- We need to show that $\mathbf{p}^T \mathbf{M}_d(y)\mathbf{p} \geq 0, \quad \forall \text{ vector: } \mathbf{p} \in \mathbb{R}^{S_{n,d}}$ (definition of PSD matrix)

$$\mathbf{p}^T \mathbf{M}_d(y)\mathbf{p} = \mathbf{p}^T \mathrm{E}[B_d(x) B_d^T(x)]\mathbf{p} = \mathrm{E}[\mathbf{p}^T B_d(x) B_d^T(x)\mathbf{p}]$$

(definition of Moment matrix)

$$= \mathrm{E}[(\mathbf{p}^T B_d(x))(\mathbf{p}^T B_d(x))^T] = \mathrm{E}[p^2(x)] = \int p^2(x) d\mu \geq 0 \Longrightarrow \mathbf{M}_d(y) \succcurlyeq 0$$

Polynomial of order d:
Coefficients vector and monomials

$$p(x) = \mathbf{p}^T B_d(x)$$

2

Moments of every measure μ in $\mathbb{R}^n$ satisfies the moment condition:

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d$$

We want to show that given a measure and its moments sequence, μ, y $\Longrightarrow$ $\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d$

- We need to show that $\mathbf{p}^T \mathbf{M}_d(y) \mathbf{p} \geq 0, \quad \forall \text{ vector: } \mathbf{p} \in \mathbb{R}^{S_{n,d}}$ (definition of PSD matrix)

$$\mathbf{p}^T \mathbf{M}_d(y) \mathbf{p} = \mathbf{p}^T \underbrace{\mathrm{E}[B_d(x) B_d^T(x)]}_{} \mathbf{p} = \mathrm{E}[\mathbf{p}^T B_d(x) B_d^T(x) \mathbf{p}]$$

(definition of Moment matrix)

$$= \mathrm{E}[\underbrace{(\mathbf{p}^T B_d(x))}_{}(\mathbf{p}^T B_d(x))^T] = \mathrm{E}[p^2(x)] = \int p^2(x) d\mu \geq 0 \Longrightarrow \mathbf{M}_d(y) \succcurlyeq 0$$

Polynomial of order d:
Coefficients vector and monomials
$p(x) = \mathbf{p}^T B_d(x)$

$$\mathbf{M}_d(y) \succcurlyeq 0 \;\equiv \mathrm{E}_\mu[p^2(x)] \geq 0, \quad \forall p(x) \in \mathbb{R}_{2d}[x]$$

3

Moments of every measure μ on the set $\mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0 \mid_{i=1}^{m}\}$ satisfies the moment condition:

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \qquad \mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d, \quad i = 1, .., m.$$

3

Moments of every measure μ on the set $\mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0 \mid_{i=1}^{m}\}$ satisfies the moment condition:

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \qquad \mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d, \quad i = 1, .., m.$$

We want to show that given measure and its moments sequence,

$$\mu, \; supp(\mu) \subset \mathbf{K}, \quad y \quad \Longrightarrow \quad \begin{array}{l} \mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \\ \mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d, \quad i = 1, .., m. \end{array}$$

- *Eq 3.15*: Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

3

Moments of every measure μ on the set $\mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0 \mid_{i=1}^{m}\}$ satisfies the moment condition:

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \qquad \mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d, \quad i = 1, .., m.$$

We want to show that given measure and its moments sequence,

$$\mu, \ supp(\mu) \subset \mathbf{K}, \quad y \quad \Longrightarrow \quad \begin{array}{c} \mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \\ \mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d, \quad i = 1, .., m. \end{array}$$

We need to show that $\mathbf{p}^T \mathbf{M}_d(g_i y) \mathbf{p} \geq 0, \quad \forall \text{ vector: } \mathbf{p} \in \mathbb{R}^{S_{n,d}}$

3

Moments of every measure μ on the set $\mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0 \;|_{i=1}^{m}\}$ satisfies the moment condition:

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \qquad \mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d, \quad i = 1, .., m.$$

We want to show that given measure and its moments sequence,

$$\mu, \; supp(\mu) \subset \mathbf{K}, \quad y \quad \Longrightarrow \quad \begin{array}{c} \mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \\ \mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d, \quad i = 1, .., m. \end{array}$$

We need to show that $\mathbf{p}^T \mathbf{M}_d(g_i y)\mathbf{p} \geq 0, \quad \forall$ vector: $\mathbf{p} \in \mathbb{R}^{S_{n,d}}$

$$\mathbf{p}^T \mathbf{M}_d(gy)\mathbf{p} = \mathbf{p}^T \underbrace{\mathrm{E}[g_i(x) B_d(x) B_d^T(x)]}\mathbf{p} = \mathrm{E}[g_i(x)\mathbf{p}^T B_d(x) B_d^T(x)\mathbf{p}]$$

(definition of Localizing matrix)

3

Moments of every measure μ on the set $\mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0 \ |_{i=1}^m\}$ satisfies the moment condition:

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \qquad \mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d, \quad i = 1, .., m.$$

We want to show that given measure and its moments sequence,

$$\mu, \ supp(\mu) \subset \mathbf{K}, \quad y \quad \Longrightarrow \quad \begin{array}{c} \mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \\ \mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d, \quad i = 1, .., m. \end{array}$$

We need to show that $\mathbf{p}^T \mathbf{M}_d(g_i y)\mathbf{p} \geq 0, \quad \forall$ vector: $\mathbf{p} \in \mathbb{R}^{S_{n,d}}$

$$\mathbf{p}^T \mathbf{M}_d(gy)\mathbf{p} = \mathbf{p}^T \mathrm{E}[g_i(x) B_d(x) B_d^T(x)]\mathbf{p} = \mathrm{E}[g_i(x)\mathbf{p}^T B_d(x) B_d^T(x)\mathbf{p}]$$

(definition of Localizing matrix)

$$= \mathrm{E}[g_i(x)(\mathbf{p}^T B_d(x))(\mathbf{p}^T B_d(x))^T] = \mathrm{E}[g_i(x)p^2(x)] = \int g_i(x)p^2(x)d\mu \geq 0 \Longrightarrow \mathbf{M}_d(g_i y) \succcurlyeq 0$$

Polynomial of order d:
Coefficients vector and monomials

$$p(x) = \mathbf{p}^T B_d(x)$$

$g_i \geq 0$ on the set $\mathbf{K}$

3

Moments of every measure μ on the set $\mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0 \mid_{i=1}^m\}$ satisfies the moment condition:

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \qquad \mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d, \quad i = 1, .., m.$$

We want to show that given measure and its moments sequence,

$$\mu,\ supp(\mu) \subset \mathbf{K}, \quad y \quad \Longrightarrow \quad \begin{array}{l} \mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \\ \mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d, \quad i = 1, .., m. \end{array}$$

We need to show that $\mathbf{p}^T \mathbf{M}_d(g_i y)\mathbf{p} \geq 0, \quad \forall$ vector: $\mathbf{p} \in \mathbb{R}^{S_{n,d}}$

$$\mathbf{p}^T \mathbf{M}_d(gy)\mathbf{p} = \mathbf{p}^T \underbrace{\mathrm{E}[g_i(x) B_d(x) B_d^T(x)]}\mathbf{p} = \mathrm{E}[g_i(x)\mathbf{p}^T B_d(x) B_d^T(x)\mathbf{p}]$$

(definition of Localizing matrix)

$$= \mathrm{E}[g_i(x)\underbrace{(\mathbf{p}^T B_d(x))}(\mathbf{p}^T B_d(x))^T] = \mathrm{E}[g_i(x)p^2(x)] = \int g_i(x)p^2(x)d\mu \geq 0 \Longrightarrow \mathbf{M}_d(g_i y) \succcurlyeq 0$$

Polynomial of order d:
Coefficients vector and monomials

$$p(x) = \mathbf{p}^T B_d(x)$$

$g_i \geq 0$ on the set $\mathbf{K}$

whenever $supp(\mu) \subset \{x \in \mathbb{R}^n : g_i(x) \geq 0\}$ $\qquad \mathbf{M}_d(g_i y) \succcurlyeq 0 \equiv \mathrm{E}_\mu[g_i p^2(x)] \geq 0, \quad \forall p(x) \in \mathbb{R}[x]$

4

Any infinite sequence y that satisfies $\mathbf{M}_d(y) \succcurlyeq 0, \mathbf{M}_d(g_i y) \succcurlyeq 0, \ \forall d, \ i = 1, .., m$ are the moment seq. of measure μ supported on compact (Archimedean) set $\mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0 \ |_{i=1}^m\}$

We want to show that

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d$$
$$\mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d, \quad i = 1, .., m. \quad \Rightarrow \quad \text{y moment sequence of measur: } \mu, \ supp(\mu) \subset \mathbf{K}$$

- *Theorems 3.1,3.8 , Chapter 3*: Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

4

Any infinite sequence y that satisfies $\mathbf{M}_d(y) \succcurlyeq 0, \mathbf{M}_d(g_i y) \succcurlyeq 0, \ \forall d, \ i = 1, .., m$ are the moment seq. of measure μ supported on compact (Archimedean) set $\mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0 \mid_{i=1}^{m}\}$

We want to show that

$$\begin{aligned} &\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \\ &\mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d, \quad i = 1, .., m. \end{aligned} \Longrightarrow \begin{aligned} &\text{y moment sequence of measur:} \\ &\mu, \ supp(\mu) \subset \mathbf{K} \end{aligned}$$

The proof relies on the following Theorem:

Theorem(Riesz-Haviland): sequence y has a measure μ on the closed set $\mathbf{K}$ if and only if

$$\mathrm{E}_\mu[p(x)] \geq 0, \quad \forall \text{ polynomilas } p(x) \geq 0 \text{ on } \mathbf{K}$$

- *Theorems 3.1,3.8 , Chapter 3*: Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

4

Any infinite sequence y that satisfies $\mathbf{M}_d(y) \succcurlyeq 0, \mathbf{M}_d(g_i y) \succcurlyeq 0, \;\forall d, \; i = 1, .., m$ are the moment seq. of measure μ supported on compact (Archimedean) set $\mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0 \mid_{i=1}^m\}$

We want to show that

$$\begin{array}{l} \mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \\ \mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d, \quad i = 1, .., m. \end{array} \Rightarrow \begin{array}{c} \text{y moment sequence of measur:} \\ \mu, \; supp(\mu) \subset \mathbf{K} \end{array}$$

The proof relies on the following Theorem:

Theorem(Riesz-Haviland): sequence y has a measure μ on the closed set $\mathbf{K}$ if and only if

$$\mathrm{E}_\mu[p(x)] \geq 0, \quad \forall \text{ polynomilas } p(x) \geq 0 \text{ on } \mathbf{K}$$

- Not that this theorem is not very practical as we do not have an explicit characterization of nonnegative Polynomials on a general closed set.

- *Theorems 3.1,3.8 , Chapter 3*: Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

4

Any infinite sequence y that satisfies $\mathbf{M}_d(y) \succcurlyeq 0, \mathbf{M}_d(g_i y) \succcurlyeq 0, \ \forall d, \ i = 1, .., m$ are the moment seq. of measure μ supported on compact (Archimedean) set $\mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0 \ |_{i=1}^m\}$

We want to show that

$$\begin{aligned} &\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \\ &\mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d, \quad i = 1, .., m. \end{aligned} \Rightarrow \begin{aligned} &\text{y moment sequence of measur:} \\ &\mu, \ supp(\mu) \subset \mathbf{K} \end{aligned}$$

The proof relies on the following Theorem:

Theorem(Riesz-Haviland): sequence y has a measure μ on the closed set $\mathbf{K}$ if and only if

$$\mathrm{E}_\mu[p(x)] \geq 0, \quad \forall \text{ polynomilas } p(x) \geq 0 \text{ on } \mathbf{K}$$

- Not that this theorem is not very practical as we do not have an explicit characterization of nonnegative Polynomials on a general closed set.

➢ Hence, we need to show that for given sequence y $\Rightarrow$ $\mathrm{E}_\mu[p(x)] \geq 0, \quad \forall \text{ polynomilas } p(x) \geq 0 \text{ on } \mathbf{K}$

- *Theorems 3.1,3.8 , Chapter 3*: Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

- Hence, we need to show that for given sequence y $\Rightarrow$ $\mathrm{E}_\mu[p(x)] \geq 0, \quad \forall \text{ polynomilas } p(x) \geq 0 \text{ on } \mathbf{K}$

- Hence, we need to show that for given sequence y $\Rightarrow$ $\mathrm{E}_\mu[p(x)] \geq 0, \quad \forall \underbrace{\text{polynomilas } p(x) \geq 0 \text{ on } \mathbf{K}}_{p(x) = \sigma_0(x) + \sum_{i=1}^m \sigma_i(x) g_i(x) \quad \sigma_0(x), \sigma_i(x) \in SOS}$

- Hence, we need to show that for given sequence y $\Rightarrow$ $\mathrm{E}_\mu[p(x)] \geq 0, \quad \forall \underbrace{\text{polynomilas } p(x) \geq 0 \text{ on } \mathbf{K}}_{p(x) = \sigma_0(x) + \sum_{i=1}^m \sigma_i(x) g_i(x) \quad \sigma_0(x), \sigma_i(x) \in SOS}$

$$\mathrm{E}_\mu[p(x)] = \mathrm{E}[\sigma_0(x) + \textstyle\sum_{i=1}^m \sigma_i(x) g_i(x)] = \mathrm{E}[\sigma_0(x)] + \sum_{i=1}^m \mathrm{E}[\sigma_i(x) g_i(x)]$$

$$= \mathrm{E}[\textstyle\sum_j f_j^2(x)] + \sum_{i=1}^m \mathrm{E}\left[\left(\sum_j h_j^2(x)\right) g_i(x)\right]$$

- Hence, we need to show that for given sequence y $\Rightarrow$ $\mathrm{E}_\mu[p(x)] \geq 0, \quad \underbrace{\forall \text{ polynomilas } p(x) \geq 0 \text{ on } \mathbf{K}}$

$$p(x) = \sigma_0(x) + \sum_{i=1}^m \sigma_i(x) g_i(x) \quad \sigma_0(x), \sigma_i(x) \in SOS$$

$$\mathrm{E}_\mu[p(x)] = \mathrm{E}[\sigma_0(x) + \textstyle\sum_{i=1}^m \sigma_i(x) g_i(x)] = \mathrm{E}[\sigma_0(x)] + \sum_{i=1}^m \mathrm{E}[\sigma_i(x) g_i(x)]$$

$$= \mathrm{E}[\overbrace{\textstyle\sum_j f_j^2(x)}] + \sum_{i=1}^m \mathrm{E}[\overbrace{\left(\textstyle\sum_j h_j^2(x)\right)} g_i(x)]$$

We know that sequence y satisfies

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \quad \equiv \mathrm{E}_\mu[p^2(x)] \geq 0, \quad \forall p(x) \in \mathbb{R}[x]$$

$$\mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d \quad \equiv \mathrm{E}_\mu[g_i p^2(x)] \geq 0, \quad \forall p(x) \in \mathbb{R}[x]$$

- Hence, we need to show that for given sequence y $\Rightarrow$ $\mathrm{E}_\mu[p(x)] \geq 0, \quad \underbrace{\forall \text{ polynomilas } p(x) \geq 0 \text{ on } \mathbf{K}}_{p(x)=\sigma_0(x)+\sum_{i=1}^m \sigma_i(x) g_i(x) \quad \sigma_0(x), \sigma_i(x) \in SOS}$

$$\mathrm{E}_\mu[p(x)] = \mathrm{E}[\sigma_0(x) + \textstyle\sum_{i=1}^m \sigma_i(x) g_i(x)] = \mathrm{E}[\sigma_0(x)] + \textstyle\sum_{i=1}^m \mathrm{E}[\sigma_i(x) g_i(x)]$$

$$= \underbrace{\mathrm{E}[\textstyle\sum_j f_j^2(x)]}_{+} + \textstyle\sum_{i=1}^m \underbrace{\mathrm{E}[\left(\textstyle\sum_j h_j^2(x)\right) g_i(x)]}_{+}$$

We know that sequence y satisfies

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \quad \equiv \mathrm{E}_\mu[p^2(x)] \geq 0, \quad \forall p(x) \in \mathbb{R}[x]$$

$$\mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d \ \equiv \mathrm{E}_\mu[g_i p^2(x)] \geq 0, \quad \forall p(x) \in \mathbb{R}[x]$$

- Hence, we need to show that for given sequence y $\Rightarrow$ $\mathrm{E}_\mu[p(x)] \geq 0, \quad \forall \underbrace{\text{polynomilas } p(x) \geq 0 \text{ on } \mathbf{K}}_{p(x) = \sigma_0(x) + \sum_{i=1}^m \sigma_i(x) g_i(x) \quad \sigma_0(x), \sigma_i(x) \in SOS}$

$$\mathrm{E}_\mu[p(x)] = \mathrm{E}[\sigma_0(x) + \sum_{i=1}^m \sigma_i(x) g_i(x)] = \mathrm{E}[\sigma_0(x)] + \sum_{i=1}^m \mathrm{E}[\sigma_i(x) g_i(x)]$$

$$= \underbrace{\mathrm{E}[\sum_j f_j^2(x)]}_{+} + \sum_{i=1}^m \underbrace{\mathrm{E}[\left(\sum_j h_j^2(x)\right) g_i(x)]}_{+}$$

We know that sequence y satisfies

$$\mathbf{M}_d(y) \succcurlyeq 0, \quad \forall d \quad \equiv \mathrm{E}_\mu[p^2(x)] \geq 0, \quad \forall p(x) \in \mathbb{R}[x]$$

$$\mathbf{M}_d(g_i y) \succcurlyeq 0, \quad \forall d \quad \equiv \mathrm{E}_\mu[g_i p^2(x)] \geq 0, \quad \forall p(x) \in \mathbb{R}[x]$$

$\Rightarrow$ $\mathrm{E}_\mu[p(x)] \geq 0$ $\xRightarrow{\text{Theorem (Riesz-Haviland)}}$ sequence y has a measure μ on the set $\mathbf{K}$

- Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

- Jean Bernard Lasserre, “Global Optimization with Polynomials and the Problem of Moments”, SIAM J. Optim., 11(3), 796–817, 2001.

- Jean Bernard Lasserre “The Moment-SOS hierarchy”, 2018

- Nie, J. “Optimality Conditions and Finite Convergence of Lasserre’s Hierarchy”, Math. Program. Ser. A 146, 97–121, 2014.

- Section 4: Monique Laurent, “Sums Of Squares, Moment Matrices and Optimization Over Polynomials”, 2010, https://homepages.cwi.nl/~monique/files/moment-ima-update-new.pdf

Gloptipoly: http://homepages.laas.fr/henrion/software/gloptipoly3/