Moments and Positive Polynomials for Optimization III:
PUTINAR VERSUS KARUSH-KUHN-TUCKER

Jean B. Lasserre

LAAS-CNRS and Institute of Mathematics, Toulouse, France

Master 2 Optimization: Paris-Orsay, 2016
Recall the GLOBAL optimization problem $P$:

$$f^* := \min_x \{ f(x) \mid g_j(x) \geq 0, j = 1, \ldots, m \},$$

where $f, g_j \in \mathbb{R}[X]$. Hence, the feasible set

$$K := \{ x \in \mathbb{R}^n \mid g_j(x) \geq 0, j = 1, \ldots, m \}$$

is a basic semi-algebraic set.
Let \( f^* := \min_x \{ f(x) : g_j(x) \geq 0, \ j = 1, \ldots, m \} \) and let \( x^* \in K \) be a minimizer at a **LOCAL** minimum.

**Karush-Kuhn-Tucker (KKT) OPTIMALITY CONDITIONS**

There exist **NONNEGATIVE SCALAR MULTIPLIERS** \( \lambda \in \mathbb{R}^m \) such that:

\[
\nabla \left[ f(x^*) - \sum_{j=1}^{m} \lambda_j g_j(x^*) \right] = 0. \quad \lambda_j g_j(x^*) = 0; \quad \lambda_j \geq 0
\]
Under some constraint qualifications:

I. The \textbf{KKT-optimality} conditions are \textit{necessary} for $x^*$ to be a \textbf{LOCAL} minimizer only.

II. If $f$ and $-g_j$ are concave, the \textbf{KKT-optimality} conditions are also \textit{sufficient} for $x^*$ to be a \textbf{GLOBAL} minimizer.
Under some constraint qualifications:

I. The KKT-optimality conditions are necessary for \( x^* \) to be a LOCAL minimizer only.

II. If \( f \) and \( -g_j \) are concave, the KKT-optimality conditions are also sufficient for \( x^* \) to be a GLOBAL minimizer.
IN GENERAL, \( x^* \) IS NOT a global minimizer of the LAGRANGIAN

\[
x \mapsto L(x) := f(x) - f^* - \sum_{j=1}^{m} \lambda_j g_j(x)
\]

but ONLY a stationary point!

However, in the CONVEX case

\( x^* \) is a global minimizer of the Lagrangian \( L \) and:

\[
L \geq 0 \quad \text{on } \mathbb{R}^n; \quad L(x^*) = 0; \quad \nabla L(x^*) = 0
\]

which provides a certificate of global optimality since

\[
L \geq 0 \quad \Rightarrow \quad f(x) - f^* = \sum_{i=1}^{m} \lambda_i g_i(x) + p(x) \geq 0
\]
IN GENERAL, \( x^* \) IS NOT a global minimizer of the LAGRANGIAN

\[
x \mapsto L(x) := f(x) - f^* - \sum_{j=1}^{m} \lambda_j g_j(x)
\]

but ONLY a stationary point!

However, in the CONVEX case

\( x^* \) is a global minimizer of the Lagrangian \( L \) and:

\[
L \geq 0 \text{ on } \mathbb{R}^n; \quad L(x^*) = 0; \quad \nabla L(x^*) = 0
\]

which provides a certificate of global optimality since

\[
L \geq 0 \quad \Rightarrow \quad f(x) - f^* = \sum_{i=1}^{m} \lambda_i g_i(x) + p(x) \geq 0
\]
IN GENERAL, $x^*$ IS NOT a global minimizer of the LAGRANGIAN

$$x \mapsto L(x) := f(x) - f^* - \sum_{j=1}^{m} \lambda_j g_j(x)$$

but ONLY a stationary point!

However, in the CONVEX case

$x^*$ is a global minimizer of the Lagrangian $L$ and:

$$L \geq 0 \text{ on } \mathbb{R}^n; \quad L(x^*) = 0; \quad \nabla L(x^*) = 0$$

which provides a certificate of global optimality since

$$L \geq 0 \quad \Rightarrow \quad f(x) - f^* = \sum_{i=1}^{m} \lambda_i g_j(x) + p(x) \geq 0.$$
Putinar’s representation theorem (Positivstellensatz)

\[ f(x) = \sigma_0(x) + \sum_{j=1}^{m} \sigma_j(x) g_j(x), \quad \forall x \in \mathbb{R}^n, \]

(for some s.o.s. polynomials \((\sigma_j)) \) ....

holds for polynomials \(f\) that are STRICTLY POSITIVE on \(K\).

However, by recent results from Marshall (2009), Nie (2012)

it also holds GENERICALLY for polynomials \(f\) that are only NONNEGATIVE on \(K\)!
Putinar’s representation theorem (Positivstellensatz)

\[ f(x) = \sigma_0(x) + \sum_{j=1}^{m} \sigma_j(x) g_j(x), \quad \forall x \in \mathbb{R}^n, \]

(for some s.o.s. polynomials \((\sigma_j))\) ....

holds for polynomials \(f\) that are STRICTLY POSITIVE on \(K\).

However, by recent results from Marshall (2009), Nie (2012)

it also holds GENERICALLY for polynomials \(f\) that are only NONNEGATIVE on \(K\)!
Theorem (Marshall, Nie)

Let $x^* \in K$ be a global minimizer of

$$P: \quad f^* = \min \{ f(x) : g_j(x) \geq 0, \ j = 1, \ldots, m \}.$$ 

and assume that:

(i) The gradients $\{\nabla g_j(x^*)\}$ are linearly independent,

(ii) Strict complementarity holds ($\lambda^*_j g_j(x^*) = 0$ for all $j$.)

(iii) Second-order sufficiency conditions hold at $(x^*, \lambda^*) \in K \times \mathbb{R}^m_+.$

Then $f(x) - f^* = \sigma^*_0(x) + \sum_{j=1}^m \sigma^*_j(x)g_j(x), \ \forall x \in \mathbb{R}^n, \text{ for some SOS polynomials } \{\sigma^*_j\}.$

Moreover, the conditions (i)-(ii)-(iii) HOLD GENERICALLY!
**Theorem (Marshall, Nie)**

Let $x^* \in K$ be a global minimizer of

$$
P : \quad f^* = \min \left\{ f(x) : g_j(x) \geq 0, \ j = 1, \ldots, m \right\}.
$$

and assume that:

(i) The gradients $\{\nabla g_j(x^*)\}$ are linearly independent,

(ii) Strict complementarity holds ($\lambda^*_j g_j(x^*) = 0$ for all $j$.)

(iii) Second-order sufficiency conditions hold at $(x^*, \lambda^*) \in K \times \mathbb{R}^m_+$.

Then $f(x) - f^* = \sigma^*_0(x) + \sum_{j=1}^{m} \sigma^*_j(x)g_j(x), \ \forall x \in \mathbb{R}^n$, for some SOS polynomials $\{\sigma^*_j\}$.

Moreover, the conditions (i)-(ii)-(iii) **HOLD GENERICALLY**!
If Putinar’s Theorem holds for $f - f^*$

(see e.g. conditions in Marshall & Nie’s theorem), then:

The SOS-hierarchy has **FINITE CONVERGENCE**!

**Proof.**

By Assumption $f - f^* = \sigma_0 + \sum_{j=1}^{m} \sigma_j g_j$ for some SOS $(\sigma_j)$ of degree bounded by some $2t$. But then as soon as $2d \geq 2t + \max \deg(g_j)$,

$$
\rho_d^* = \max_{\lambda, \sigma_j} \{ \lambda : f - \lambda = \sigma_0 + \sum_{j=1}^{m} \sigma_j g_j; \deg\sigma_j g_j \leq 2d \} \geq f^*.
$$

Combining with $\rho_d^* \leq f^*$ for all $d$, yields the desired result. □
Finite convergence of the SOS hierarchy

If Putinar’s Theorem holds for $f - f^*$

(see e.g. conditions in Marshall & Nie’s theorem), then:

The SOS-hierarchy has FINITE CONVERGENCE!

Proof.

By Assumption $f - f^* = \sigma_0 + \sum_{j=1}^{m} \sigma_j g_j$ for some SOS $(\sigma_j)$ of degree bounded by some $2t$. But then as soon as $2d \geq 2t + \max \deg(g_j)$,

$$\rho_d^* = \max_{\lambda, \sigma_j} \{ \lambda : f - \lambda = \sigma_0 + \sum_{j=1}^{m} \sigma_j g_j; \ \deg \sigma_j g_j \leq 2d \} \geq f^*.$$

Combining with $\rho_d^* \leq f^*$ for all $d$, yields the desired result. □
Hence **GENERICALLY**, solving the **global polynomial optimization problem**

\[
\inf_x \{ f(x) : x \in K \},
\]

where \( K \) is the compact set \( \{ x : g_j(x) \geq 0, j = 1, \ldots, m \} \),

... reduces to solving the **semidefinite program**

\[
\sup_{\lambda, \sigma_j} \{ \lambda : f - \lambda = \sigma_0 + \sum_{j=1}^{m} \sigma_j g_j; \ \deg \sigma_j g_j \leq 2d \},
\]

but where \( d \) is not known in advance!
Hence **GENERICALLY**, solving the **global polynomial optimization problem**

\[
\inf_{x} \{ f(x) : x \in K \},
\]

where \( K \) is the compact set \( \{ x : g_j(x) \geq 0, j = 1, \ldots, m \} \),

... reduces to solving the **semidefinite program**

\[
\sup_{\lambda, \sigma_j} \{ \lambda : f - \lambda = \sigma_0 + \sum_{j=1}^{m} \sigma_j g_j; \quad \deg \sigma_j g_j \leq 2d \},
\]

but where \( d \) is not known in advance!
If Putinar’s Theorem holds for $f - f^*$ (only $\geq 0$ on $K$), then

the EXTENDED LAGRANGIAN polynomial

$$x \mapsto \Psi(x) := f(x) - f^* - \sum_{j=1}^{m} \sigma_j(x) g_j(x) \quad (= \sigma_0(x))$$

(with s.o.s. MULTIPLIERS $\sigma_j \in \mathbb{R}[X]$ instead of scalar $\lambda \in \mathbb{R}^m$)

is s.o.s.! (hence $\Psi \geq 0$ on $\mathbb{R}^n$), and satisfies:

$$\nabla \Psi(x^*) = \nabla f(x^*) - \sum_{j=1}^{m} \left[ \sigma_j(x^*) \nabla g_j(x^*) \right] = 0$$

$$\sigma_j(x^*) g_j(x^*) = 0 \quad \forall j \quad \text{(and so } \Psi(x^*) = 0)$$

$\rightarrow (x^*, \lambda^*)$ is a KKT pair.
If Putinar’s Theorem holds for $f - f^*$ (only $\geq 0$ on $K$), then

the EXTENDED LAGRANGIAN polynomial

$$x \mapsto \Psi(x) := f(x) - f^* - \sum_{j=1}^{m} \sigma_j(x) g_j(x) \quad (= \sigma_0(x))$$

(with s.o.s. MULTIPLIERS $\sigma_j \in \mathbb{R}[X]$ instead of scalar $\lambda \in \mathbb{R}^m$)

is s.o.s.! (hence $\Psi \geq 0$ on $\mathbb{R}^n$), and satisfies:

$$\nabla \Psi(x^*) = \nabla f(x^*) - \sum_{j=1}^{m} \sigma_j(x^*) \nabla g_j(x^*) = 0$$

$$= \underbrace{\nabla f(x^*)}_{\lambda^*_j \geq 0} - \sum_{j=1}^{m} \underbrace{\sigma_j(x^*) \nabla g_j(x^*)}_{\lambda^*_j \geq 0}$$

$$\sigma_j(x^*) g_j(x^*) = 0 \quad \forall j \quad \text{(and so } \Psi(x^*) = 0)$$

$$(x^*, \lambda^*) \text{ is a KKT pair.}$$
If Putinar’s Theorem holds for $f - f^*$ (only $\geq 0$ on $K$), then the EXTENDED LAGRANGIAN polynomial

$$x \mapsto \Psi(x) := f(x) - f^* - \sum_{j=1}^{m} \sigma_j(x) g_j(x) \quad (= \sigma_0(x))$$

(with s.o.s. MULTIPLIERS $\sigma_j \in \mathbb{R}[X]$ instead of scalar $\lambda \in \mathbb{R}^m$) is s.o.s.! (hence $\Psi \geq 0$ on $\mathbb{R}^n$), and satisfies:

$$\nabla \Psi(x^*) = \nabla f(x^*) - \sum_{j=1}^{m} \underbrace{\nabla \sigma_j(x^*) \nabla g_j(x^*)}_{\lambda_j^* \geq 0} = 0$$

$$\sigma_j(x^*) g_j(x^*) = 0 \quad \forall j \quad \text{(and so } \Psi(x^*) = 0)$$

$\rightarrow (x^*, \lambda^*)$ is a KKT pair.
That is ...

\( x^* \) is a **GLOBAL MINIMIZER**
of the **EXTENDED LAGRANGIAN** \( \psi \) on \( \mathbb{R}^n \)!

So when Putinar's representation holds

for the polynomial \( f - f^* \) (which is only nonnegative on \( K \))

it provides a **global optimality certificate** for \( f^* \) and \( x^* \in K \)

... the analogue

in **nonconvex polynomial optimization**
of the **KKT-optimality conditions** for the general **convex case**

....

a highly nontrivial extension ..!!!
That is ...

\[ x^* \text{ is a GLOBAL MINIMIZER of the EXTENDED LAGRANGIAN } \psi \text{ on } \mathbb{R}^n! \]

So when Putinar's representation holds for the polynomial \( f - f^* \) (which is only nonnegative on \( K \))

it provides a **global optimality certificate** for \( f^* \) and \( x^* \in K \)

... the analogue in nonconvex polynomial optimization

of the **KKT-optimality conditions** for the general convex case

... a highly nontrivial extension ..!!!
That is ...

\[ x^* \text{ is a GLOBAL MINIMIZER} \]
\[ \text{of the EXTENDED LAGRANGIAN } \Psi \text{ on } \mathbb{R}^n! \]

So when Putinar’s representation holds

for the polynomial \( f - f^* \) (which is only nonnegative on \( K \))

it provides a **global optimality certificate** for \( f^* \) and \( x^* \in K \)

... the analogue

in **nonconvex polynomial optimization**

of the **KKT-optimality conditions** for the general **convex case**

....

a **highly nontrivial extension** ..!!!
An important property

On non active constraints

Let \((x^*, \lambda) \in K \times \mathbb{R}_+^m\) be a KKT point with \(x^*\) a global minimizer of \(P\) and suppose that the constraint \(g_j \geq 0\) is not active at \(x^*\), i.e., \(g_j(x^*) > 0\).

Then,

in contrast to KKT optimality conditions where the associated scalar multiplier \(\lambda_j\) vanishes \((\lambda_j = 0)\), ...

the s.o.s. “multiplier” \(\sigma_j\) of the extended Lagrangian \(\Psi\) does not vanish in general, but ... \(\sigma_j(x^*) = 0 (= \lambda_j)!\)
An important property

On non active constraints

Let \((x^*, \lambda) \in \mathbf{K} \times \mathbb{R}^m_+\) be a KKT point with \(x^*\) a global minimizer of \(P\) and suppose that the constraint \(g_j \geq 0\) is not active at \(x^*\), i.e., \(g_j(x^*) > 0\).

Then,

in contrast to KKT optimality conditions where the associated scalar multiplier \(\lambda_j\) vanishes (\(\lambda_j = 0\)), ...  

the s.o.s. “multiplier” \(\sigma_j\) of the extended Lagrangian \(\Psi\) does not vanish in general, but ... \(\sigma_j(x^*) = 0 (= \lambda_j)\)!
Indeed, even if NOT ACTIVE at $x^*$,

IN THE NONCONVEX case, the constraint $g_j(x) \geq 0$ MAY STILL BE IMPORTANT because if deleted, the global optimum $f^*$ may strictly decrease to $\theta < f^*$.

Therefore in the case where $\theta < f^*$

the constraint $g_j(x) \geq 0$ MUST PLAY a ROLE in Putinar’s representation of the polynomial $f - f^*$, i.e., its associated s.o.s. weight $\sigma_j$ is NOT trivial.

Otherwise if $\sigma_j = 0$, i.e., if $f - f^* = \sigma_0 + \sum_{k \neq j} \sigma_k g_k$ then

$$\theta = \min_{x} \{ f(x) : g_k(x) \geq 0, \forall k \neq j \} = f^*.$$

However, its VALUE at $x^*$ VANISHES ($\sigma_j(x^*) = 0$)!
Indeed, even if NOT ACTIVE at $x^*$,

IN THE NONCONVEX case, the constraint $g_j(x) \geq 0$ MAY STILL BE IMPORTANT because if deleted, the global optimum $f^*$ may strictly decrease to $\theta < f^*$.

Therefore in the case where $\theta < f^*$

the constraint $g_j(x) \geq 0$ MUST PLAY a ROLE in Putinar’s representation of the polynomial $f - f^*$, i.e., its associated s.o.s. weight $\sigma_j$ is NOT trivial.

Otherwise if $\sigma_j = 0$, i.e., if $f - f^* = \sigma_0 + \sum_{k \neq j} \sigma_k g_k$ then

\[
\theta = \min_x \{ f(x) : g_k(x) \geq 0, \forall k \neq j \} = f^*.
\]

However, its VALUE at $x^*$ VANISHES ($\sigma_j(x^*) = 0$)!
Indeed, even if NOT ACTIVE at $x^*$,

IN THE NONCONVEX case, the constraint $g_j(x) \geq 0$ MAY STILL BE IMPORTANT because if deleted, the global optimum $f^*$ may strictly decrease to $\theta < f^*$.

Therefore in the case where $\theta < f^*$

the constraint $g_j(x) \geq 0$ MUST PLAY a ROLE in Putinar’s representation of the polynomial $f - f^*$, i.e., its associated s.o.s. weight $\sigma_j$ is NOT trivial.

Otherwise if $\sigma_j = 0$, i.e., if $f - f^* = \sigma_0 + \sum_{k \neq j} \sigma_k g_k$ then

$$\theta = \min_{x} \{ f(x) : g_k(x) \geq 0, \forall k \neq j \} = f^*.$$

However, its VALUE at $x^*$ VANISHES ($\sigma_j(x^*) = 0$)!
Example: Consider the one-dimensional problem

\[ \mathbf{P} : \quad f^* = \min_x \{-x \mid x^2 = 1; \ 1/2 - x \geq 0\}, \]

with \( X \mapsto g_1(X) = X^2 - 1 \) and \( X \mapsto g_2(X) := 0.5 - X \).

\( x^* = -1 \) is a global minimizer with global minimum \( f^* = 1 \).

\( (x^*, \lambda) = (-1, (1/2, 0)) \) is a KKT pair, and \( \lambda_2 = 0 \) because the constraint \( g_2(x) \geq 0 \) is not active at \( x^* = -1 \).

Of course, \( x^* \) is not a global minimum of the Lagrangian

\[ f - \lambda_1 g_1 - \lambda_2 g_2 = -X - 1/2(X^2 - 1) = -X^2/2 - X + 1/2. \]
Example: Consider the one-dimensional problem

$$
\mathbf{P} : \quad f^* = \min_x \{-x \mid x^2 = 1; \ 1/2 - x \geq 0\},
$$

with $X \mapsto g_1(X) = X^2 - 1$ and $X \mapsto g_2(X) := 0.5 - X$.

$x^* = -1$ is a global minimizer with global minimum $f^* = 1$.

$(x^*, \lambda) = (-1, (1/2, 0))$ is a KKT pair, and $\lambda_2 = 0$ because the constraint $g_2(x) \geq 0$ is not active at $x^* = -1$.

Of course, $x^*$ is not a global minimum of the Lagrangian $f - \lambda_1 g_1 - \lambda_2 g_2 = -X - 1/2(X^2 - 1) = -X^2/2 - X + 1/2$. 
Example: Consider the one-dimensional problem

\[ P: \quad f^* = \min_x \{-x \mid x^2 = 1; \ 1/2 - x \geq 0\}, \]

with \( X \mapsto g_1(X) = X^2 - 1 \) and \( X \mapsto g_2(X) := 0.5 - X \).

\( x^* = -1 \) is a global minimizer with global minimum \( f^* = 1 \).

\((x^*, \lambda) = (-1, (1/2, 0))\) is a KKT pair, and \( \lambda_2 = 0 \) because the constraint \( g_2(x) \geq 0 \) is not active at \( x^* = -1 \).

Of course, \( x^* \) is not a global minimum of the Lagrangian

\[ f - \lambda_1 g_1 - \lambda_2 g_2 = -X - 1/2(X^2 - 1) = -X^2/2 - X + 1/2. \]
Example: Consider the one-dimensional problem

\[ P : \quad f^* = \min_x \{-x \mid x^2 = 1; \ 1/2 - x \geq 0\}, \]

with \( X \mapsto g_1(X) = X^2 - 1 \) and \( X \mapsto g_2(X) := 0.5 - X \).

\( x^* = -1 \) is a global minimizer with global minimum \( f^* = 1 \).

\((x^*, \lambda) = (-1, (1/2, 0))\) is a KKT pair, and \( \lambda_2 = 0 \) because the constraint \( g_2(x) \geq 0 \) is not active at \( x^* = -1 \).

Of course, \( x^* \) is not a global minimum of the Lagrangian

\[ f - \lambda_1 g_1 - \lambda_2 g_2 = -X - 1/2(X^2 - 1) = -X^2/2 - X + 1/2.\]
But we also have Putinar’s representation

\[ f - f^* = -X - 1 = (X + 3/2)(X^2 - 1) + (X + 1)^2(1/2 - X). \]

The s.o.s. (polynomial) multiplier \( x \mapsto \sigma_2(X) := (X + 1)^2 \)
vanishes at \( x^* = -1 \), also a global minimizer of the Lagrangian

\[ f - \sigma_1 g_1 - \sigma_2 g_2 = -X - (X + 3/2)(X^2 - 1) - (X + 1)^2(1/2 - X) \]

(here constant \( \equiv 1 \)).

Even if not active at \( x^* \), the constraint \( g_2(x) \geq 0 \) is important
because if deleted, \( f^* \rightarrow -1 < 1 \). Therefore, it MUST have a
nontrivial s.o.s. multiplier in the representation of \( f - f^* \).
But we also have Putinar’s representation
\[ f - f^* = -X - 1 = (X + 3/2)(X^2 - 1) + (X + 1)^2(1/2 - X). \]

The s.o.s. (polynomial) multiplier \( x \mapsto \sigma_2(X) := (X + 1)^2 \)
vanishes at \( x^* = -1 \), also a global minimizer of the Lagrangian
\[ f - \sigma_1 g_1 - \sigma_2 g_2 = -X - (X + 3/2)(X^2 - 1) - (X + 1)^2(1/2 - X) \]
(here constant \( \equiv 1 \)).

Even if not active at \( x^* \), the constraint \( g_2(x) \geq 0 \) is important
because if deleted, \( f^* \rightarrow -1 < 1 \). Therefore, it MUST have a
nontrivial s.o.s. multiplier in the representation of \( f - f^* \).
But we also have Putinar’s representation

\[ f - f^* = -X - 1 = (X + 3/2)(X^2 - 1) + (X + 1)^2(1/2 - X). \]

The s.o.s. (polynomial) multiplier \( \sigma_2(X) := (X + 1)^2 \)
vanishes at \( x^* = -1 \), also a global minimizer of the Lagrangian

\[ f - \sigma_1 g_1 - \sigma_2 g_2 = -X - (X + 3/2)(X^2 - 1) - (X + 1)^2(1/2 - X) \]

(here constant \( \equiv 1 \)).

Even if not active at \( x^* \), the constraint \( g_2(x) \geq 0 \) is important because if deleted, \( f^* \rightarrow -1 < 1 \). Therefore, it MUST have a nontrivial s.o.s. multiplier in the representation of \( f - f^* \).
An important observation

The \textbf{MOMENT-SOS} approach is a \textbf{GENERAL PURPOSE} method

\textbf{AIMING AT SOLVING NP-hard PROBLEMS}

\textbf{... and ANY GENERAL PURPOSE approach} should have the \textbf{HIGHLY DESIRABLE} feature to behave efficiently for problems considered “\textbf{EASY}!”

Otherwise \textbf{... would you buy such a package?}
An important observation

The **MOMENT-SOS** approach is a **GENERAL PURPOSE** method aiming at solving NP-hard problems.

... and any **GENERAL PURPOSE** approach should have the **HIGHLY DESIRABLE** feature to behave efficiently for problems considered “**EASY**”!

Otherwise ... would you buy such a package?
An important observation

The **MOMENT-SOS** approach is a **GENERAL PURPOSE** method

**AIMING AT SOLVING** NP-hard **PROBLEMS**

... and ANY GENERAL PURPOSE approach should have the **HIGHLY DESIRABLE** feature to behave efficiently for problems considered "**EASY"!

Otherwise ... would you buy such a package?
Optimization problems

\[ f^* = \min \left\{ f(x) : g_j(x) \geq 0, \quad j = 1, \ldots, m \right\} \]

where \( f \) and \( -g_j \) are convex,

are considered \textbf{EASY} as they can be solved efficiently by appropriate methods (e.g. using logarithmic barrier method).

A polynomial \( f \in \mathbb{R}[X] \) is \textbf{SOS-CONVEX} if its Hessian \( \nabla^2 f(x) \) factors as \( L(x) L(x)^T \) for some matrix polynomial \( L \in \mathbb{R}[X]^{n \times p} \) (for some \( p \)).
Optimization problems

\[ f^* = \min \{ f(x) : g_j(x) \geq 0, \quad j = 1, \ldots, m \} \]

where \( f \) and \(-g_j\) are convex,

are considered EASY as they can be solved efficiently by appropriate methods (e.g. using logarithmic barrier method).

A polynomial \( f \in \mathbb{R}[X] \) is SOS-CONVEX if its Hessian \( \nabla^2 f(x) \) factors as \( L(x) L(x)^T \) for some matrix polynomial \( L \in \mathbb{R}[X]^{n \times p} \) (for some \( p \)).
Optimization problems

\[ f^* = \min \{ f(x) : g_j(x) \geq 0, \ j = 1, \ldots, m \} \]

where \( f \) and \(-g_j\) are convex,

are considered EASY as they can be solved efficiently by appropriate methods (e.g. using logarithmic barrier method).

A polynomial \( f \in \mathbb{R}[X] \) is SOS-CONVEX if its Hessian \( \nabla^2 f(x) \) factors as \( L(x) L(x)^T \) for some matrix polynomial \( L \in \mathbb{R}[X]^{n \times p} \) (for some \( p \)).
A remarkable property of the SOS hierarchy: I

When solving the optimization problem

\[ \textbf{P} : \quad f^* = \min \{ f(x) : g_j(x) \geq 0, \; j = 1, \ldots, m \} \]

one does NOT distinguish between **CONVEX**, **CONTINUOUS NON CONVEX**, and **0/1 (and DISCRETE)** problems! A boolean variable \( x_i \) is modelled via the equality constraint “\( x_i^2 - x_i = 0 \).

In Non Linear Programming (NLP),

modeling a 0/1 variable with the polynomial equality constraint “\( x_i^2 - x_i = 0 \)

and applying a standard descent algorithm would be considered “stupid”!

Each class of problems has its own *ad hoc* tailored algorithms.
A remarkable property of the SOS hierarchy: 

When solving the optimization problem

\[ P : \quad f^* = \min \{ f(x) : g_j(x) \geq 0, \ j = 1, \ldots, m \} \]

one does NOT distinguish between **CONVEX**, **CONTINUOUS NON CONVEX**, and 0/1 (and **DISCRETE**) problems! A boolean variable \( x_i \) is modelled via the equality constraint “\( x_i^2 - x_i = 0 \)”. 

In Non Linear Programming (NLP),

modeling a 0/1 variable with the polynomial equality constraint “\( x_i^2 - x_i = 0 \)” and applying a standard descent algorithm would be considered “stupid”!

Each class of problems has its own **ad hoc** tailored algorithms.
A remarkable property of the SOS hierarchy: I

When solving the optimization problem

\[ \mathbf{P} : \quad f^* = \min \{ f(x) : g_j(x) \geq 0, \quad j = 1, \ldots, m \} \]

one does NOT distinguish between CONVEX, CONTINUOUS NON CONVEX, and 0/1 (and DISCRETE) problems! A boolean variable \( x_i \) is modelled via the equality constraint “\( x_i^2 - x_i = 0 \).”

In Non Linear Programming (NLP),

modeling a 0/1 variable with the polynomial equality constraint “\( x_i^2 - x_i = 0 \)"

and applying a standard descent algorithm would be considered “stupid”!

Each class of problems has its own \emph{ad hoc} tailored algorithms.
Even though the moment-SOS approach **DOES NOT SPECIALIZE** to each class of problems:

- It recognizes the class of (easy) SOS-convex problems as **FINITE CONVERGENCE** occurs at the **FIRST** relaxation in the hierarchy.
- Finite convergence also occurs for general convex problems and generically for non convex problems → (NOT true for the LP-hierarchy.)
- The **SOS-hierarchy** dominates other lift-and-project hierarchies (i.e. provides the best lower bounds) for hard 0/1 combinatorial optimization problems! The Computer Science community talks about a **META-Algorithm**.
Even though the moment-SOS approach **DOES NOT SPECIALIZE** to each class of problems:

- It **recognizes** the class of (easy) **SOS-convex problems** as **FINITE CONVERGENCE** occurs at the **FIRST** relaxation in the hierarchy.
- Finite convergence also occurs for general convex problems and generically for non convex problems → (NOT true for the LP-hierarchy.)
- The **SOS-hierarchy** dominates other **lift-and-project** hierarchies (i.e. provides the best lower bounds) for hard 0/1 combinatorial optimization problems! The Computer Science community talks about a **META-Algorithm**.
Even though the moment-SOS approach **DOES NOT SPECIALIZE** to each class of problems:

- It recognizes the class of (easy) **SOS-convex problems** as **FINITE CONVERGENCE** occurs at the **FIRST** relaxation in the hierarchy.
- Finite convergence also occurs for general convex problems and generically for non convex problems → (NOT true for the LP-hierarchy.)
- The **SOS-hierarchy** dominates other lift-and-project hierarchies (i.e. provides the best lower bounds) for hard 0/1 combinatorial optimization problems! The Computer Science community talks about a **META-Algorithm**.
Even though the moment-SOS approach DOES NOT SPECIALIZE to each class of problems:

- It recognizes the class of (easy) SOS-convex problems as **FINITE CONVERGENCE** occurs at the FIRST relaxation in the hierarchy.
- Finite convergence also occurs for general convex problems and generically for non convex problems
- → (NOT true for the LP-hierarchy.)
- The **SOS-hierarchy** dominates other lift-and-project hierarchies (i.e. provides the best lower bounds) for hard 0/1 combinatorial optimization problems! The Computer Science community talks about a **META-Algorithm**.
Even though the moment-SOS approach **DOES NOT SPECIALIZE** to each class of problems:

- It **recognizes** the class of (easy) **SOS-convex problems** as **FINITE CONVERGENCE** occurs at the **FIRST** relaxation in the hierarchy.
- Finite convergence also occurs for general convex problems and generically for non convex problems
- $\rightarrow$ (NOT true for the **LP-hierarchy**.)
- The **SOS-hierarchy** dominates other **lift-and-project** hierarchies (i.e. provides the best lower bounds) for hard 0/1 combinatorial optimization problems! The Computer Science community talks about a **META-Algorithm**.
A remarkable property: II

FINITE CONVERGENCE of the SOS-hierarchy is GENERIC!

... and provides a GLOBAL OPTIMALITY CERTIFICATE,

the analogue for the NON CONVEX CASE of the KKT-OPTIMALITY conditions in the CONVEX CASE!