LECTURE 5

Recall that we have

(SAO)

$$f^* = \operatorname{Inf} f(x)$$

s.t. $x \in S(g)$

$$\left(S(g) = \{ x \in \mathbb{R}^n : \underset{i=1..m}{g_i(x) \geq 0} \} \right)$$

NON-CONVEX

FINITE-DIMENSIONAL $S(g) \subset \mathbb{R}^n$

$$=$$

(MP)

$$m^* = \operatorname{Inf} \int_{S(g)} f \, d\mu$$

s.t. $\operatorname{supp}(\mu) \subset S(g)$

$$\int_{S(g)} d\mu = 1$$

LINEAR!

but INFINITE-DIMENSIONAL

Indeed, the space of Borel measures has infinite dimension, as a $\mathbb{R}$-vector space. [Exercise]

When $S(g) = \mathbb{R}^n$ (global optimization) we have characterized

$$\mathbb{R}[x]_{2d}^{\vee} \supset \mathcal{P}(\mathbb{R}^n)_{2d}^{*} = \text{conical hull}\Big(\{ \underset{\text{point evaluations}}{l_v : \mathbb{R}[x]_{2d} \to \mathbb{R}} \} \Big) =$$

$$= \{ y = (y_\alpha)_{\mathbb{N}^n_{2d}} : y \text{ has a representing measure } \mu \}$$

$$(y_\alpha = \int x^\alpha d\mu \ \ \forall \alpha \in \mathbb{N}^n_{2d})$$

$$\left[f = \sum c_\alpha x^\alpha : \mathcal{L}_y(f) = \sum c_\alpha y_\alpha \right]$$

$$\mathbb{R}[x]_{2d}^{\vee} \hookrightarrow \Sigma_{n,2d}^{*} = \{ l \in \mathbb{R}[x]_{2d}^{\vee} : Q_l \succeq 0 \}$$

$$Q_l(f) := l(f^2)$$

$$\mathcal{P}(\mathbb{R}^n)_{2d}^{*} \subset \Sigma_{n,2d}^{*}$$

↑ we will re-name this matrix.

Remark

Another proof that $\mathcal{P}(\mathbb{R}^n)^*_{2d} \subset \Sigma^*_{n,2d}$:

$v \in \mathbb{R}^n$, $l_v \in \mathcal{P}(\mathbb{R}^n)^*_{2d}$

$$Q_{l_v}(f) = l_v(f^2) = f^2(v) \geq 0 \quad \forall f \in \mathbb{R}[x]_d$$

$\rightarrow l_v \in \Sigma^*_{n,2d}$.

This means that for l to belong to the cone $\mathcal{P}(\mathbb{R}^n)^*_{2d}$, it is necessary (but not sufficient!!) that $l(f^2)$ is positive semidefinite.

Example $n=1$, $2d=4$. We consider the sequence $y=(y_0,y_1,y_2,y_3,y_4)=(1,1,1,1,2)$.

We get the functional $\mathcal{L}_y$ that acts:

$$g = g_0 + g_1 x + \dots + g_4 x^4 \curvearrowright \mathcal{L}_y(g) = g_0 y_0 + \dots + g_4 y_4 =$$
$$= g_0 + g_1 + g_2 + g_3 + 2 g_4 .$$

We prove that $y \sim \mathcal{L}_y \in \Sigma^*_{1,4}$. This is true because in this case

$$Q_{\mathcal{L}_y} \sim \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix} \succeq 0$$

we test on ↙ $g = f^2$.

Let's prove explicitely that

$$\forall f \in \mathbb{R}[x]_2 \rightarrow \mathcal{L}_y(f^2) \geq 0.$$

We ~~the~~ can suppose w.l.o.g. that f is monic:

$$f = f_0 + f_1 x + x^2$$

$$\rightarrow \mathcal{L}_y(f^2) = \mathcal{L}_y\big(f_0^2 + f_1^2 x^2 + x^4 + 2 f_0 f_1 x + + 2 f_1 x^3 + 2 f_0 x^2 \big) =$$

$$= \mathcal{L}_y\big(f_0^2 + 2 f_0 f_1 x + (f_1^2 + 2 f_0) x^2 + 2 f_1 x^3 + x^4 \big) =$$

$$= f_0^2 + 2 f_0 f_1 + f_1^2 + 2 f_0 + 2 f_1 + 2 =$$

$$= (f_0 + 1)^2 + (f_1 + 1)^2 + 2 f_0 f_1 =$$

Rename $X = f_0 + 1$, $Y = f_1 + 1$

$$= X^2 + Y^2 + 2(X-1)(Y-1) =$$

$$= (X+Y)^2 - 2(X+Y) + 2 =$$

Rename $Z = X + Y$

$$= Z^2 - 2Z + 2 > 0 \quad \text{because } \Delta = -4 < 0.$$

However, the sequence y has no representing measure μ, hence $\mathcal{L}_y \in \Sigma^*_{1,4} \setminus \mathcal{P}(\mathbb{R})^*_4$.
(no proof)

Now we give a meaning to the matrix Q_ℓ.

Moment and Localizing matrices

Let $y = (y_\alpha)_{\alpha \in \mathbb{N}^n_{2d}} \subset \mathbb{R}$, ~~let~~ with $d \in \mathbb{N}$.

The moment matrix in degree d of y is the symmetric reel matrix of size $\binom{n+d}{d} \times \binom{n+d}{d}$ $M_d(y)$ given by:

$$M_d(y)(\alpha, \beta) = \mathcal{L}_y(x^\alpha x^\beta) = y_{\alpha+\beta}$$

with $\alpha, \beta \in \mathbb{N}^n_d$.

The rows and columns of $M_d(y)$ are indexed by the elements of the basis $\langle x^\alpha \rangle = \mathbb{R}[x]_d$.

For instance, for $n=2$, $d=2$:

$$M_2(y) = \left[\begin{array}{c|cc|ccc} y_{00} & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ \hline y_{10} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ \hline y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{array}\right] \swarrow$$

Block Hankel (or generalized Hankel)

For $n=1$, we get Hankel matrices:

$y = (y_0\ y_1\ y_2\ y_3\ \cdots\ y_{2d})$

$$\to M_d(y) = \begin{pmatrix} y_0 & y_1 & y_2 & \iddots & \\ y_1 & y_2 & \iddots & & \\ y_2 & \iddots & & & \\ & & & & y_{2d-1} \\ & & & y_{2d-1} & y_{2d} \end{pmatrix}$$

$M_d(y)$ defines the following form

$$\langle \cdot , \cdot \rangle_y : \mathbb{R}[x]_d \times \mathbb{R}[x]_d \longrightarrow \mathbb{R}$$

$$\langle p, q \rangle_y = \mathcal{L}_y(pq) = \langle \underline{p}, M_d(y)\underline{q} \rangle$$

$$\Big\downarrow$$

$$\underline{p}^T M_d(y)\, \underline{q}$$

where $\underline{p}$, $\underline{q}$ are the column vectors of coefficients of p and q, in the monomial basis.

Exercise For $y = (1\ 1\ 1\ 1\ 2)$ as above, write explicitely $M_2(y)$ and the quadratic form $\langle \cdot , \cdot \rangle_y$.

Remark If y has a representing measure μ then, for every $f \in \mathbb{R}[x]_d$

$$\langle p, p \rangle_y = \mathcal{L}_y(p^2) = \int_{\mathbb{R}^n} p^2 \, d\mu \geq 0$$

because p^2 is positive.

$$\| \quad \langle \underline{p}, M_d(y)\underline{p} \rangle .$$

Hence $M_d(y) \succeq 0$!!

Exercise ~~Con~~ Let $\mu = \delta_{x_0}$ be the Dirac delta with support on a fixed $x_0 \in \mathbb{R}$. Write down

$17.1 + 14.2 + 13 \approx 45\, m^2$ + balcone

its sequence y of moments.

Localizing matrix: $y = (y_\alpha)_{\alpha \in \mathbb{N}^n_d}$

Given $g \in \mathbb{R}[x]$, with coefficients g_γ, $\gamma \in \mathbb{N}^n$ we define the <u>localizing matrix</u> of g:

$$M_d(gy) := \Big(\mathcal{L}_y\big(g(x)\, x^\alpha x^\beta\big)\Big)_{\alpha,\beta}$$

$$\mathcal{L}_y\big(g(x)\, x^\alpha x^\beta\big) = \sum_{\gamma \in \mathbb{N}^n} g_\gamma\, y_{\gamma+\alpha+\beta} \qquad \forall\, \alpha, \beta \in \mathbb{N}^n_d .$$

It means that:

- → to construct $M_d(y)$, we apply the Riesz functional $\mathcal{L}_y$ entry-wise to the matrix $v_d(x)\, v_d(x)^T$;
- → to construct $M_d(gy)$, we do the same to the matrix $g(x)\, v_d(x)\, v_d(x)^T$.

Remember that we defined matrices B_α, C_α:

$$v_d(x)\, v_d(x)^T = \sum_{\alpha \in \mathbb{N}^n_{2d}} B_\alpha x^\alpha$$

$d' = \deg g$

$$: \quad g(x)\, v_d(x)\, v_d(x)^T = \sum_{\alpha \in \mathbb{N}^n_{2d+d'}} C_\alpha x^\alpha$$

Donc on a

$$\left.\begin{aligned} M_d(y) &= \sum_{\alpha\in\mathbb{N}^n_{2d}} B_\alpha y_\alpha \\ M_d(gy) &= \sum_{\alpha\in\mathbb{N}^n_{2d+d'}} C_\alpha y_\alpha \end{aligned}\right\} \text{Linearizations of } \underset{v_d(x)\,v_d(x)^T}{\underset{\shortparallel}{\sum B_\alpha x^\alpha}},\ \sum C_\alpha x^\alpha$$

For the localizing matrices too:

$$\langle p, q\rangle_{gy} = \mathcal{L}_y(g\,p\,q) =$$

bilinear form $\langle \cdot , \cdot \rangle_{gy} : \mathbb{R}[x]_d \times \mathbb{R}[x]_d \to \mathbb{R}$.

If $p \in \mathbb{R}[x]_d$

$$\langle p, p\rangle_{gy} = \mathcal{L}_y(g\,p^2) \overset{\substack{\text{If } y \text{ has a representing measure}\\ \downarrow}}{=} \int_{\mathbb{R}^n} g\,p^2\,d\mu$$

If $\operatorname{supp}\mu \subset S(g) = \{x\in\mathbb{R}^n : g(x) \geqslant 0\}$

then $\langle p, p\rangle_{gy} \geqslant 0 \quad \forall p$, that is

$$M_d(gy) \succcurlyeq 0 \ !!$$

Recall The Riesz functional has the property to "linearize polynomials": $p = \sum p_\alpha x^\alpha$

$$\mathcal{L}_y(p) = \sum p_\alpha \mathcal{L}_y(x^\alpha) = \sum p_\alpha y_\alpha =$$

If y has a rep. measure μ

$$\swarrow$$

$$= \sum p_\alpha \int_{\mathbb{R}^n} x^\alpha d\mu = \int_{\mathbb{R}^n} \sum p_\alpha x^\alpha d\mu = \int_{\mathbb{R}^n} p \, d\mu .$$

<u>Exercise</u> (To do now) $n=2$, $g = 1+2x_1+3x_2$.

What is $M_1(gy)$?

Here $y = (y_\alpha)_{\alpha \in \mathbb{N}^2} = (y_{00}, y_{10}, y_{01}, y_{20}, y_{11}, y_{02} \cdots)$

But $d=1$, so we consider the truncated sequence up to order $2d + \deg(g) = 3$.

$$y = (y_{00}, y_{10}, y_{01}, \cdots, y_{03})$$

By definition:

$$M_1(gy) = \begin{pmatrix} \overset{\mathcal{L}_y(g x^{\binom{0}{0}} x^{\binom{0}{0}})}{\overset{\shortparallel}{1+2y_{10}+3y_{01}}} & \mathcal{L}_y(g x^{\binom{1}{0}} x^{\binom{0}{0}}) & \cdots \\ \mathcal{L}_y(g x^{\binom{0}{0}} x^{\binom{1}{0}}) & \ddots & \\ \vdots & & \end{pmatrix} =$$

$$= \begin{array}{c|ccc} & 1 & x_1 & x_2 \\ \hline 1 & 1+2y_{10}+3y_{01} & y_{10}+2y_{20}+3y_{11} & y_{01}+2y_{11}+3y_{02} \\ x_1 & y_{10}+2y_{20}+3y_{11} & y_{20}+2y_{30}+3y_{21} & y_{11}+2y_{21}+3y_{12} \\ x_2 & y_{01}+2y_{11}+3y_{02} & y_{11}+2y_{21}+3y_{12} & y_{02}+2y_{12}+3y_{03} \end{array}$$

13h30

Flat extensions

Let $d \in \mathbb{N}$, and $y = (y_\alpha)_{\alpha \in \mathbb{N}^n_{2d}} \in \mathbb{R}^{\binom{n+2d}{n}}$ be a real sequence.

We say that $M_d(y)$ is a flat extension of $M_{d-1}(y)$ if $\operatorname{rk} M_d(y) = \operatorname{rk} M_{d-1}(y)$.

Theorem (Curto-Fialkow, 1996)

If $M_d(y) \succeq 0$ and $M_d(y)$ is a flat extension of $M_{d-1}(y)$, then y has a representing measure.

Example $y = (2,1,1,1,1)$. Then $M_2(y)$ is a flat extension of $M_1(y)$:

$$M_2(y) = \left[\begin{array}{cc|c} \multicolumn{2}{c}{M_1(y)} & \\ 2 & 1 & 1 \\ 1 & 1 & 1 \\ \hline 1 & 1 & 1 \end{array}\right] \qquad \operatorname{rk} M_1(y) = \operatorname{rk} M_2(y) = 2$$

Definition (Atomic measure) An atom of a measure μ is a measurable set E such that:

1) $\mu(E) > 0$
2) $\forall\, A \subset E$ with $\mu(A) < \mu(E)$ then $\mu(A) = 0$.

A measure μ is *atomic* if there exists an atom for μ. It is *r-atomic* if it contains exactly r atoms.

For instance

1. δ_x Dirac deltas are atomic, as their combinations $\sum c_i \delta_{x_i}$

2. The Lebesgue measure λ_n on $\mathbb{R}^n$ is non-atomic

3. X = finite set; σ-algebra on X given by the power-set of X (the set of all subsets). Then every set $\{x\}$ of one element is an atom.

Theorem y admits an r-atomic representing measure (with $r = \operatorname{rk} M_d(y)$) if and only if $M_d(y)$ admits a flat extension $M_{d+1}(\tilde{y})$, with $\tilde{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n_{2(d+1)}}$.

Q What can we say about $y = (2,1,1)$? (in the previous example).

Putinar's Theorem revisited

$$S(g) = \{x \in \mathbb{R}^n : g_i(x) \geq 0 \ \forall i = 1..m\}$$

We suppose that the quadratic module

$$(g_0 = 1) \qquad Q(g) = \{\sum_{i=0}^{m} \sigma_i g_i : \sigma_i \text{ are sums of squares}\}$$

is archimedean.

(For instance, if $S(g)$ is compact and $S(g) \subset B(0,R)$, $R > 0$, then we suppose $u(x) = R^2 - \sum x_i^2 \in Q(g)$).

Theorem (Putinar for measures)

$Q(g)$ archimedean. Then a sequence $y = (y_\alpha)_{\alpha \in \mathbb{N}^n}$ has a representing measure with support in $S(g)$ if and only if

$$M_d(y) \succeq 0 \quad \forall d \in \mathbb{N}$$

and $M_d(g_i y) \succeq 0 \quad \forall d \in \mathbb{N}, \ \forall i = 1 \dots m$.

We use the compact notation $M(y)$, $M(g\,y)$ for the infinite-size matrices

$$M(y) = \begin{pmatrix} M_d(y) & & \\ & M_{d+1}(y) & \\ & & \ddots \end{pmatrix}$$

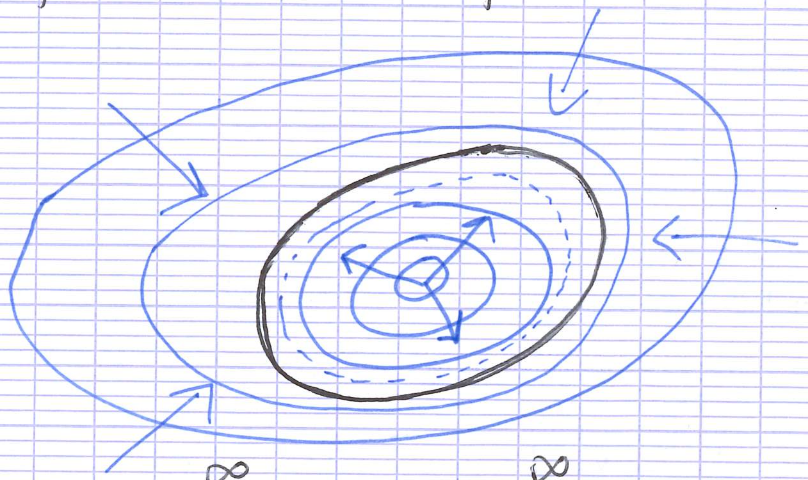
and we say

$$M(y) \succeq 0 \text{ if } M_d(y) \succeq 0 \ \forall d \in \mathbb{N}.$$

as in Putinar's theorem.

Approximation of $\mathcal{P}(S(g))$

Given a set $C \subset V$, we aim at "approximating" C from "inside" and from "outside":



$$C = \bigcup_{i=1}^{\infty} C_i^I = \bigcap_{i=1}^{\infty} C_i^O \quad \text{and}$$

C_i^I, C_i^O have better properties ("tractable" sets, for us they will be spectrahedra, feasible sets of SDP).

Truncated preordering

$$P_k(g) = \{ \sum_{J \subset \{1..m\}} \sigma_J g_J : \sigma_J \in \Sigma_n, \deg(\sigma_J g_J) \le 2k \}$$

Truncated Quadratic Module

$Q_k(g) = \left\{ \sum_{j=0}^{m} \sigma_j g_j \; : \; \sigma_j \in \Sigma_n , \; \deg(\sigma_j g_j) \leq 2k \right\}$

$P_k(g)$, $Q_k(g)$ convex cones in $\mathbb{R}[x]_{2k} \simeq \mathbb{R}^{\binom{2k+n}{n}}$

Theorem (Inner SD-approximation)

(1) If $S(g)$ is compact, then

$$\mathcal{P}(S(g))_d = \overline{\bigcup_{k=0}^{\infty} P_k(g) \cap \mathbb{R}[x]_d}$$

(2) If $Q(g)$ is archimedean, then

$$\mathcal{P}(S(g))_d = \overline{\bigcup_{k=0}^{\infty} Q_k(g) \cap \mathbb{R}[x]_d} .$$

Proof

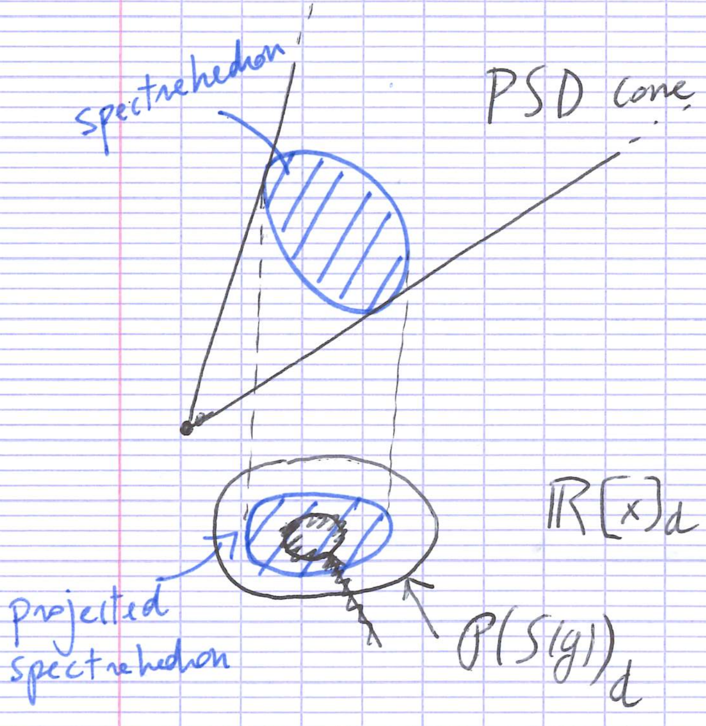
$\mathcal{P}(S(g))_d$ is closed and $P_k(g) \cap \mathbb{R}[x]_d \subset \mathcal{P}(S(g))_d$

hence $\overline{\bigcup_{k=0}^{\infty} P_k(g) \cap \mathbb{R}[x]_d} \subset \mathcal{P}(S(g))_d$.

Then, if $f \in \mathcal{P}(S(g))_d$, then $f + \frac{1}{\ell} > 0$ on $S(g)$. By Schmudgen positivstellensatz $f + \frac{1}{\ell} \in P(g) = \bigcup_{k=0}^{\infty} P_k(g)$, hence $\exists\, k_\ell \in \mathbb{N}$ such that $f + \frac{1}{\ell} \in P_{k_\ell}(g)$, $\forall\, k \geq k_\ell$.

By letting $l \to \infty$ we have $f \in \overline{\bigcup_{k=0}^{\infty} P_k(g) \cap \mathbb{R}[x]_d}$.
In the second case, we apply Putinar's PSS. $\square$

Important $P_k^{(g)} \cap \mathbb{R}[x]_d$ and $Q_k(g) \cap \mathbb{R}[x]_d$ are tractable sets because they are the projections of feasible sets of SD problems (projection of spectrahedra, the sections of the cone of PSD matrices).
(⟹ LECTURE 3).



So $P(S(g))_d$ admits ← an inner approximation by projected spectrahedra.

The outer approximation is easier, given by spectrahedra (with no projection).

Define the Borel measure

$$E \mapsto \mu(E) = \int_E \exp\left(-\sum_{i=1}^{n} |x_i|\right) d\varphi$$

where φ is a Borel measure with support
$\operatorname{supp} \varphi = S(g)$.

Let $y_\alpha = \int_{\mathbb{R}^n} x^\alpha d\mu$ for all $\alpha \in \mathbb{N}^n$.

Theorem (Outer SD-approximation)

~~Let~~ The set

$$\mathcal{P}(S(g))_d^k = \{ f \in \mathbb{R}[x]_d : M_k(fy) \succcurlyeq 0 \}$$

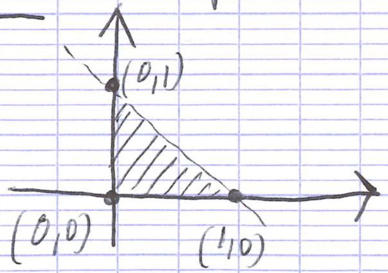
is a spectrahedron (section of the PSD cone).

Moreover $\mathcal{P}(S(g))_d^k \supset \mathcal{P}(S(g))_d^{k+1} \supset \dots \supset \mathcal{P}(S(g))_d$ $\forall k$, and

$$\mathcal{P}(S(g))_d = \bigcap_{k=0}^{\infty} \mathcal{P}(S(g))_d^k .$$

Example (Simplex in $\mathbb{R}^2$)

$S(g) =$

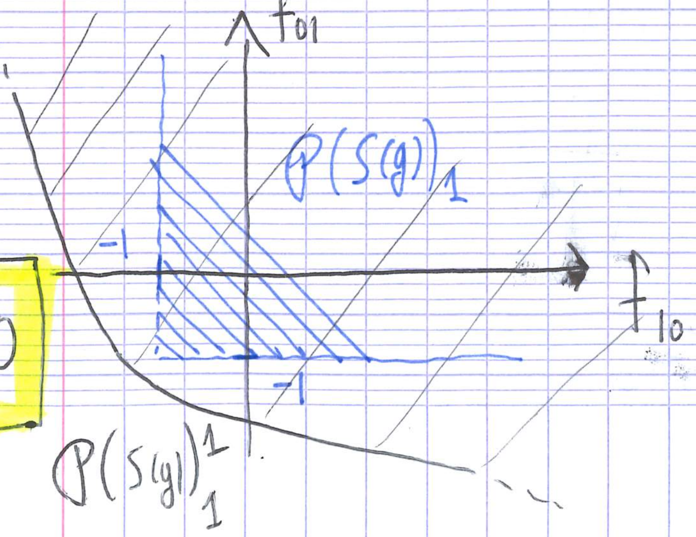


Then we know that

$$\mathcal{P}(S(g))_1 = \{ f_{00} + f_{10}x_1 + f_{01}x_2 \in \mathbb{R}[x]_1 : \begin{array}{l} f_{00} \geqslant 0 \\ f_{00} + f_{10} \geqslant 0 \\ f_{00} + f_{01} \geqslant 0 \end{array} \}$$

Supp $f_{00} = 1$:

15h00



outer approx.

$$\mathcal{P}(S(g))_1^1 = \{ f : M_1(fy) \succcurlyeq 0 \}$$

Recap on representation of positive polynomials

	Univariate Polynomials	Multivariate Polynomials	
		• $d=2$, • $(n,2d)=(2,4)$ Hilbert's Theorem: $\mathcal{P}(\mathbb{R}^n)_{2d}=\Sigma_{n,2d}$	Other cases $\mathcal{P}(\mathbb{R}^n)_{2d} \supsetneq \Sigma_{n,2d}$ SOS of rational functions!
Global case $S(g)=\mathbb{R}^n$	$f\geq 0$ $\mathcal{P}(\mathbb{R})_{2d}=\Sigma_{1,2d}$ $f=f_1^2+\dots+f_r^2$		
Local case $S(g)\subsetneq\mathbb{R}^n$	$I=[0,+\infty)$ or $I=[-1,1]$ Representation theorems (Polya-Szegö)	$S(g)$ = compact Schmüdgen's Positivstellensatz (exponential representation)	$P(g)$ = archimedean Putinar's PSS. (linear representation)

Exercise Minimize $f(x)=x^2+1$ over $\mathbb{R}$, solving a single SDP.

Sup of λ such that

$$\underset{\displaystyle x^2+1-\lambda}{\underset{\shortparallel}{f-\lambda}} = \begin{pmatrix}1 & x\end{pmatrix}\begin{pmatrix}x_{11} & x_{12}\\ x_{12} & x_{22}\end{pmatrix}\begin{pmatrix}1\\ x\end{pmatrix} = x_{11}+2x_{12}x+x_{22}x^2$$

$$\begin{pmatrix}x_{11} & x_{12}\\ x_{12} & x_{22}\end{pmatrix}\succeq 0$$

$$\begin{pmatrix}1-\lambda & 0\\ 0 & 1\end{pmatrix}\succeq 0$$

$x_{22}=1$

$x_{12}=0$

$x_{11}=1-\lambda$

$$\sup \lambda \quad \text{s.t.} \begin{pmatrix}1-\lambda & 0\\ 0 & 1\end{pmatrix}\succeq 0$$