On optimization on the Euclidean unit sphere

Jean B. Lasserre*

LAAS-CNRS and Institute of Mathematics, Toulouse, France

Optimization without Borders
Sotchi, July 12th - 17th, 2021
Nesterov's 60th and Protasov's 50th birthdate

⋆ Research funded by ANITI institute, within the framework of the French Program "Investing for the Future C PIA3"" under grant ANR-19-PIA3-0004

Based on :

Homogeneous polynomials and spurious local minima on the unit sphere
arXiv:2010.07066

Optimization on the Euclidean unit sphere
hal-03284041

- Homogeneous optimization on $\mathbb{S}^{n-1}$

  ☞ Simple and complete characterization of first and second-order optimality conditions

- Sparse optimization on $\mathbb{S}^{n-1}$

  ☞ (when the criterion to minimize is a function of a few linear forms)

I. Homogeneous polynomials & spurious local minima on the unit sphere

Let $\mathbb{S}^{n-1}$ be the Euclidean unit sphere $\{\mathbf{x} : \|\mathbf{x}\| \leq 1\} \subset \mathbb{R}^n$, and $\mathscr{E}_n$ the Euclidean unit ball $\{\mathbf{x} : \|\mathbf{x}\| \leq 1\}$.

Consider the optimization problem

$$\mathbf{P}: \qquad f^* = \min \{ f(\mathbf{x}) : \mathbf{x} \in \mathbb{S}^{n-1} \}$$

where f is a homogeneous polynomial of degree t.

☞ Problem **P** is NP-hard in general with several important applications, and in particular for some well-known combinatorial optimization problems.

- Finding the maximal cardinality $\alpha(G)$ of a stable set in a graph G reduces to minimizing a cubic form on $\mathbb{S}^{n-1}$.
- Deciding convexity of an n-variate form reduces to minimizing a form on $\mathbb{S}^{n-1}$.
- Deciding nonnegativity of an even-degree form reduces to minimizing this form on $\mathbb{S}^{n-1}$.
- Deciding copositivity of a symmetric matrix reduces to check whether some associated quartic form is nonnegative on $\mathbb{S}^{n-1}$.
- In quantum information, the Best Separable State problem also relates to homogeneous polynomial optimization on $\mathbb{S}^{n-1}$ (Fawzi et al.)

In the first part of the talk, f is a form of degree $t > 2$.

This talk

A complete & simple characterization of Second-Order Necessary Optimality conditions (SONC) at $\mathbf{x}^* \in \mathbb{S}^{n-1}$ solely in terms of

- $f(\mathbf{x}^*)$
- The first two smallest eigenvalues $\lambda_1(\nabla^2 f(\mathbf{x}^*))$ and $\lambda_2(\nabla^2 f(\mathbf{x}^*))$ of the Hessian $\nabla^2 f$ at $\mathbf{x}^* \in \mathbb{S}^{n-1}$.

First-Order Necessary Optimality Conditions (FONC)

A point $\mathbf{x}^* \in \mathbb{S}^{n-1}$ satisfies **(FONC)** if and only if

$$\nabla f(\mathbf{x}^*) \,=\, t\, f(\mathbf{x}^*)\,\mathbf{x}^*$$

Equivalently, if and only if

$$\|\nabla f(\mathbf{x}^*)\|^2 \,=\, t^2\, f(\mathbf{x}^*)^2\,.$$

Démonstration.

KKT-optimality conditions yield

$$\nabla f(\mathbf{x}^*) \,=\, 2\,\lambda^*\mathbf{x}^*\,,\quad \text{for some } \lambda^*$$

Euler's identity for homogeneous functions yields

$$t\,f(\mathbf{x}^*) \,=\, \langle \mathbf{x}^*, \nabla f(\mathbf{x}^*)\rangle \,=\, 2\,\lambda^*\,\|\mathbf{x}^*\|^2 \,=\, 2\,\lambda^*\,.$$

$$\Rightarrow \nabla f(\mathbf{x}^*) \,=\, t\,f(\mathbf{x}^*)\,\mathbf{x}^* \quad \text{and} \quad \|\nabla f(\mathbf{x}^*)\|^2 \,=\, t^2\,f(\mathbf{x}^*)^2\,.$$

□

Démonstration.

(Continued) Conversely, suppose that $\|\nabla f(\mathbf{x}^*)\|^2 = t^2 f(\mathbf{x}^*)^2$. Then

$$\begin{aligned}\|\nabla f(\mathbf{x}^*) - t f(\mathbf{x}^*)\,\mathbf{x}^*\|^2 &= \|\nabla f(\mathbf{x}^*)\|^2 - \underbrace{2t f(\mathbf{x}^*)\langle \nabla f(\mathbf{x}^*), \mathbf{x}^*\rangle}_{=-2t^2 f(\mathbf{x}^*)^2} \\ &\quad + t^2 f(\mathbf{x}^*)^2 \|\mathbf{x}^*\|^2 \\ &= \|\nabla f(\mathbf{x}^*)\|^2 - t^2 f(\mathbf{x}^*)^2 = 0 \\ \Rightarrow \nabla f(\mathbf{x}^*) &= t f(\mathbf{x}^*)\,\mathbf{x}^*,\end{aligned}$$

that is, **(FONC)** holds with $2\lambda^* = t f(\mathbf{x}^*)$. □

Given $\mathbf{x} \in \mathbb{R}^n$, let

$$\mathbf{x}^{\perp} := \{ \mathbf{u} \in \mathbb{S}^{n-1} : \mathbf{u} \perp \mathbf{x}^* \} .$$

Second-Order Necessary Optimality Conditions (SONC)

Definition : A point $\mathbf{x}^* \in \mathbb{S}^{n-1}$ which satisfies **(FONC)** satisfies **(SONC)** if

$$\langle \mathbf{u}, (\nabla^2 f(\mathbf{x}^*) - 2\,\lambda^*\,\mathbf{I})\,\mathbf{u} \rangle \geq 0 , \quad \forall \mathbf{u} \perp \mathbf{x}^* .$$

Equivalently, if and only if

$$\langle \mathbf{u}, \nabla^2 f(\mathbf{x}^*)\,\mathbf{u} \rangle \geq 2\,\lambda^* (= t\,f(\mathbf{x}^*)) , \quad \forall \mathbf{u} \in (\mathbf{x}^*)^{\perp} .$$

Lemma

If $\mathbf{x}^* \in \mathbb{S}^{n-1}$ *satisfies* **(FONC)** *then* $\mathbf{x}^*$ *is an eigenvector of* $\nabla^2 f(\mathbf{x}^*)$ *with associated eigenvalue* $t\,(t-1)\,f(\mathbf{x}^*)$.

Démonstration.

Observe that $\mathbf{x} \mapsto \partial f(\mathbf{x})/\partial x_i$ is homogeneous of degree $t-1$, for every $i = 1, \ldots, n$, and therefore by Euler's identity

$$\nabla^2 f(\mathbf{x})\,\mathbf{x} \;=\; (t-1)\,\nabla f(\mathbf{x})\,, \quad \forall \mathbf{x} \in \mathbb{R}^n\,.$$

So if $\mathbf{x}^* \in \mathbb{S}^{n-1}$ satisfies **(FONC)** then

$$\nabla^2 f(\mathbf{x}^*)\,\mathbf{x}^* \;=\; (t-1)\nabla f(\mathbf{x}^*) \;=\; t\,(t-1)\,f(\mathbf{x}^*)\,\mathbf{x}^*.$$

□

Let

$$\tau(\mathbf{x}^*) \;:=\; \min\{\langle \mathbf{u}, (\nabla^2 f(\mathbf{x}^*)\,\mathbf{u}\rangle :\; \mathbf{u} \in (\mathbf{x}^*)^\perp\,\},$$

so that if $\mathbf{x}^* \in \mathbb{S}^{n-1}$ then

$$(\textbf{SONC}) \;\Leftrightarrow \tau(\mathbf{x}^*) \;\geq\; t f(\mathbf{x}^*)\,.$$

Lemma

If $\mathbf{x}^* \in \mathbb{S}^{n-1}$ *satisfies* **(FONC)** *then*

$$\lambda_1(\nabla^2 f(\mathbf{x}^*)) \;=\; \min\,[\,t\,(t-1)\,f(\mathbf{x}^*)\,, \tau(\mathbf{x}^*)\,]\,.$$

Démonstration.

Observe that any $\mathbf{z} \in \mathbb{S}^{n-1}$ reads

$$\mathbf{z} = \alpha\,\mathbf{x}^* \oplus \beta\,\mathbf{u} :\; \mathbf{u} \in (\mathbf{x}^*)^\perp\,;\; \alpha^2 + \beta^2 = 1$$

and observe that if $\mathbf{x}^* \in \mathbb{S}^{n-1}$ satisfies **(FONC)**

$$\langle \mathbf{u}\,, \nabla^2 f(\mathbf{x}^*)\,\mathbf{x}^*\rangle \;=\; t\,(t-1)\,f(\mathbf{x}^*)\,\langle \mathbf{u}\,, \mathbf{x}^*\rangle \;=\; 0\,.$$

□

Démonstration.

Hence

$$\begin{aligned}\lambda_1(\nabla^2 f(\mathbf{x}^*)) &= \min\{\langle \mathbf{z}, \nabla^2 f(\mathbf{x}^*)\,\mathbf{z}\rangle : \mathbf{z}\in\mathbb{S}^{n-1}\}\\ &= \min\{\langle(\alpha\mathbf{x}^*+\beta\mathbf{u}), \nabla^2 f(\mathbf{x}^*)\,(\alpha\mathbf{x}^*+\beta\mathbf{u})\rangle : \alpha^2+\beta^2=1\}\\ &= \min\{\alpha^2\, t\,(t-1) f(\mathbf{x}^*)+\beta^2\,\tau(\mathbf{x}^*) : \alpha^2+\beta^2=1\}\\ &= \min\,[\,t\,(t-1)f(\mathbf{x}^*)\,,\,\tau(\mathbf{x}^*)\,]\,.\end{aligned}$$

□

In particular

If $\mathbf{x}^* \in \mathbb{S}^{n-1}$ satisfies **(FONC)** then :

$$\lambda_1(\nabla^2 f(\mathbf{x}^*)) = t\,(t-1) f(\mathbf{x}^*) \Rightarrow \tau(\mathbf{x}^*) = \lambda_2(\nabla^2 f(\mathbf{x}^*))$$

and **(SONC)** $\Leftrightarrow \quad \lambda_2(\nabla^2 f(\mathbf{x}^*)) \geq t f(\mathbf{x}^*) = \dfrac{\lambda_1(\nabla^2 f(\mathbf{x}^*))}{t-1}\,.$

Theorem

Let $\mathbf{x}^* \in \mathbb{S}^{n-1}$ *satisfy* **(FONC)** . *Then* $\mathbf{x}^*$ *satisfies* **(SONC)** *if and only if*

$$\begin{aligned} \lambda_1(\nabla^2 f(\mathbf{x}^*)) &\geq t f(\mathbf{x}^*) \quad \text{if } f(\mathbf{x}^*) \geq 0 \\ \lambda_2(\nabla^2 f(\mathbf{x}^*)) &\geq t f(\mathbf{x}^*) \quad \text{if } f(\mathbf{x}^*) < 0. \end{aligned}$$

So all **(SONC)** *points* $\mathbf{x}^* \in \mathbb{S}^{n-1}$ *are characterized by :*
- the value $f(\mathbf{x}^*)$, *and*
- The first and second smallest eigenvalues of $\nabla^2 f$ *at* $\mathbf{x}^*$.

In particular

If $\mathbf{x}^* \in \mathbb{S}^{n-1}$ satisfies **(FONC)** and $f(\mathbf{x}^*) < 0$ then :

$$\textbf{(SONC)} \Leftrightarrow \lambda_2(\nabla^2 f(\mathbf{x}^*)) \geq \frac{\lambda_1(\nabla^2 f(\mathbf{x}^*))}{t-1}$$

and $\mathbf{x}^*$ is the eigenvector associated with $\lambda_1(\nabla^2 f(\mathbf{x}^*))$.

☞ $\lambda_2(\nabla^2 f(\mathbf{x}^*))$ is significantly larger than $\lambda_1(\nabla^2 f(\mathbf{x}^*))$.

The case of odd degree forms

The only interesting local minima are negative since if $\mathbf{x} \in \mathbb{S}^{n-1}$ with $f(\mathbf{x}) > 0$ then $\mathbf{u} := -\mathbf{x} \in \mathbb{S}^{n-1}$ with value $f(\mathbf{u}) = -f(\mathbf{x}) < 0$.

So in this case

a negative local minimizer $\mathbf{x}^* \in \mathbb{S}^{n-1}$ is necessarily the eigenvector of the Hessian $\nabla^2 f(\mathbf{x}^*)$ associated with the SMALLEST EIGENVALUE $\lambda_1(\nabla^2 f(\mathbf{x}^*)) = t\,(t-1)\,f(\mathbf{x}^*)$.

Moreover

all negative local minima of f on $\mathbb{S}^{n-1}$ are all negative local minima of f on the Euclidean ball $\mathscr{E}_n$.

It is then much easier to minimize f on the convex set $\mathscr{E}_n$ rather than on the "nasty" non-convex set $\mathbb{S}^{n-1}$.

The case of odd degree forms

The only interesting local minima are negative since if $\mathbf{x} \in \mathbb{S}^{n-1}$ with $f(\mathbf{x}) > 0$ then $\mathbf{u} := -\mathbf{x} \in \mathbb{S}^{n-1}$ with value $f(\mathbf{u}) = -f(\mathbf{x}) < 0$.

So in this case

a negative local minimizer $\mathbf{x}^* \in \mathbb{S}^{n-1}$ is necessarily the eigenvector of the Hessian $\nabla^2 f(\mathbf{x}^*)$ associated with the SMALLEST EIGENVALUE $\lambda_1(\nabla^2 f(\mathbf{x}^*)) = t\,(t-1)\,f(\mathbf{x}^*)$.

Moreover

all negative local minima of f on $\mathbb{S}^{n-1}$ are all negative local minima of f on the Euclidean ball $\mathscr{E}_n$.

It is then much easier to minimize f on the convex set $\mathscr{E}_n$ rather than on the “nasty" non-convex set $\mathbb{S}^{n-1}$.

The case of cubic forms

☞ For instance if $t = 3$ (i.e., f is a cubic form) then

- $\nabla^2 f(\mathbf{x}^*)$ is a matrix whose entries are linear forms, and so
- $\mathbf{x} \mapsto \lambda_1(\nabla^2 f(\mathbf{x}))$ is a CONCAVE function
- $\mathbf{x} \mapsto \lambda_1(\nabla^2 f(\mathbf{x})) + \lambda_2(\nabla^2 f(\mathbf{x}))$ is also a CONCAVE function.

Hence the (SONC) condition

$$\lambda_2(\nabla^2 f(\mathbf{x}^*)) \geq \frac{\lambda_1(\nabla^2 f(\mathbf{x}^*))}{t-1}$$

is equivalent to the condition

$$\lambda_1(\nabla^2 f(\mathbf{x}^*)) + \lambda_2(\nabla^2 f(\mathbf{x}^*)) \geq \frac{3\,\lambda_1(\nabla^2 f(\mathbf{x}^*))}{2}$$

Theorem

An odd degree-t form f (with $t > 1$) has no spurious negative local minimum if the system

$$\begin{aligned} \nabla^2 f(\mathbf{x})\,\mathbf{x} &= t\,(t-1)\,f(\mathbf{x})\,\mathbf{x} \\ \lambda_2(\nabla^2 f(\mathbf{x})) &\geq \frac{\lambda_1(\nabla^2 f(\mathbf{x}))}{t-1} \end{aligned}$$

has a unique solution $\mathbf{x}$ with $f(\mathbf{x}) < 0$

One may also relate the critical points of f on $\mathbb{S}^{n-1}$ with the sub-varieties of a certain gradient ideal.

Let $f \in \mathbb{R}[\mathbf{x}]$ be homogeneous of degree t, and define $g \in \mathbb{R}[\mathbf{x}]$ by :

$$\mathbf{x} \mapsto g(\mathbf{x}) := f(\mathbf{x}) \, (1 - \frac{t \, \|\mathbf{x}\|^2}{t+2}) \,, \quad \mathbf{x} \in \mathbb{R}^n \,.$$

Proposition

$$\text{On } \mathbb{S}^{n-1} : \quad \nabla g(\mathbf{x}) = 0 \quad \Leftrightarrow \quad \nabla f(\mathbf{x}) = t\, f(\mathbf{x})\, \mathbf{x}$$

and so all **(FONC)** points of f are critical points of g and conversely.

$$\begin{aligned} V_{grad}(g) &:= \{\mathbf{z} \in \mathbb{C}^n : \nabla g(\mathbf{x}) = 0\} \\ &= W_0 \cup W_1 \ldots \cup W_p \end{aligned}$$

where each W_i is an irreducible subvariety and g is constant on each W_i.

☞ Hence f has no spurious local minimum on $\mathbb{S}^{n-1}$ if all the **(SONC)** points belong to a single $W_j \cap \mathbb{S}^{n-1}$ for some index j^* .

☞ an algebraic characterization

II. Sparse optimization on $\mathbb{S}^{n-1}$

Consider the problem

$$\begin{aligned}\mathbf{P}: \qquad f^* &= \min\,\{h(\mathbf{x})\,:\,\mathbf{x}\in\mathbb{S}^{n-1}\,\}\\ \text{where } h(\mathbf{x}) &= f((\ell_1\cdot\mathbf{x}),(\ell_2\cdot\mathbf{x}),\ldots,(\ell_m\cdot\mathbf{x}))\end{aligned}$$

- for some m linear forms $\ell_1,\ldots,\ell_m:\mathbb{R}^n\to\mathbb{R}$, and
- some function $f:\mathbb{R}^m\to\mathbb{R}$.

☞ If $m \ll n$ then **P** can be view a sparse optimization problem as in h, the coupling of variables only occurs through the m linear forms $(\ell_j)_{j\in[m]}$.

☞ Sometimes h is said to be *low-rank*

☞ However the constraint $\mathbf{x}\in\mathbb{S}^{n-1}$ *is not expressed* through the ℓ_j's.

Take home message

Problem **P** is in fact equivalent to an m variables problem on the Euclidean unit ball $\mathscr{E}_m := \{\mathbf{x} \in \mathbb{R}^m : \|\mathbf{x}\| \leq 1\}$ of $\mathbb{R}^m$.

- Let $\boldsymbol{\ell} \in \mathbb{R}^{m\times n}$ be the real matrix with rows $(\ell_j)_{j\in[m]}$.
- $\mathcal{L} := \boldsymbol{\ell}\,\boldsymbol{\ell}^T \in \mathbb{R}^{m\times m}$, and $\mathbf{L} := \mathcal{L}^{1/2}$,

and consider the problem

$$\mathbf{Q}: \qquad \min_{\mathbf{y}} \{ f((\mathbf{L}_1 \cdot \mathbf{y}), \ldots, (\mathbf{L}_m \cdot \mathbf{y})) : \; \mathbf{y} \in \mathscr{E}_m \} .$$

Theorem

Assume that the ℓ_j are linearly independent.
(i) Let $\mathbf{x}^* \in \mathbb{S}^{n-1}$ satisfy **(SONC)** *for problem* $\mathbf{P}$*. Then there exists $\mathbf{y}^* \in \mathscr{E}_m$ which satisfies* **(SONC)** *for problem* $\mathbf{Q}$ *and with same value*

$$f((\mathbf{L}_1 \cdot \mathbf{y}^*), \ldots, (\mathbf{L}_m \cdot \mathbf{y}^*)) = f(\boldsymbol{\ell}\,\mathbf{x}^*) = h(\mathbf{x}^*) .$$

(ii) Conversely, let $\mathbf{y}^* \in \mathscr{E}_m$ satisfy **(SONC)** *for problem* $\mathbf{Q}$*. Then there exists $\mathbf{x}^* \in \mathbb{S}^{n-1}$ which satisfies* **(SONC)** *for problem* $\mathbf{P}$ *and with same value*

$$h(\mathbf{x}^*) = f(\boldsymbol{\ell}\,\mathbf{x}^*) = f((\mathbf{L}_1 \cdot \mathbf{y}^*), \ldots, (\mathbf{L}_m \cdot \mathbf{y}^*)) .$$

- Let $\boldsymbol{\ell} \in \mathbb{R}^{m\times n}$ be the real matrix with rows $(\ell_j)_{j\in[m]}$.
- $\mathcal{L} := \boldsymbol{\ell}\,\boldsymbol{\ell}^T \in \mathbb{R}^{m\times m}$, and $\mathbf{L} := \mathcal{L}^{1/2}$,

and consider the problem

$$\mathbf{Q}: \qquad \min_{\mathbf{y}} \{ f((\mathbf{L}_1 \cdot \mathbf{y}), \ldots, (\mathbf{L}_m \cdot \mathbf{y})) : \; \mathbf{y} \in \mathscr{E}_m \} .$$

Theorem

Assume that the $\boldsymbol{\ell}_j$ are linearly independent.
(i) Let $\mathbf{x}^* \in \mathbb{S}^{n-1}$ satisfy **(SONC)** *for problem* **P**. *Then there exists $\mathbf{y}^* \in \mathscr{E}_m$ which satisfies* **(SONC)** *for problem* **Q** *and with same value*

$$f((\mathbf{L}_1 \cdot \mathbf{y}^*), \ldots, (\mathbf{L}_m \cdot \mathbf{y}^*)) \,=\, f(\boldsymbol{\ell}\,\mathbf{x}^*) \,=\, h(\mathbf{x}^*) .$$

(ii) Conversely, let $\mathbf{y}^* \in \mathscr{E}_m$ satisfy **(SONC)** *for problem* **Q**. *Then there exists $\mathbf{x}^* \in \mathbb{S}^{n-1}$ which satisfies* **(SONC)** *for problem* **P** *and with same value*

$$h(\mathbf{x}^*) \,=\, f(\boldsymbol{\ell}\,\mathbf{x}^*) \,=\, f((\mathbf{L}_1 \cdot \mathbf{y}^*), \ldots, (\mathbf{L}_m \cdot \mathbf{y}^*)) .$$

Key observation

The KKT-optimality conditions at $\mathbf{x}^* \in \mathbb{S}^{n-1}$ read :

$$\boldsymbol{\ell}^T \nabla f((\boldsymbol{\ell}_1 \cdot \mathbf{x}^*), \ldots, (\boldsymbol{\ell}_m \cdot \mathbf{x}^*)) \,=\, 2\,\lambda^* \mathbf{x}^*\,,$$

for some λ^*.

In other words, whenever $\lambda^* \neq 0$

any candidate local minimizer $\mathbf{x}^* \in \mathbb{S}^{n-1}$ necessarily satisfies

$$\mathbf{x}^* \,\in\, \mathrm{Span}(\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_m)$$

☞ It is enough to search in $\mathrm{Span}(\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_m) \cap \mathbb{S}^{n-1}$!

Key observation

The KKT-optimality conditions at $\mathbf{x}^* \in \mathbb{S}^{n-1}$ read :

$$\boldsymbol{\ell}^T \nabla f((\boldsymbol{\ell}_1 \cdot \mathbf{x}^*), \ldots, (\boldsymbol{\ell}_m \cdot \mathbf{x}^*)) \,=\, 2\,\lambda^* \mathbf{x}^*\,,$$

for some λ^*.

In other words, whenever $\lambda^* \neq 0$

any candidate local minimizer $\mathbf{x}^* \in \mathbb{S}^{n-1}$ necessarily satisfies

$$\mathbf{x}^* \,\in\, \mathrm{Span}(\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_m)$$

☞ It is enough to search in $\mathrm{Span}(\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_m) \cap \mathbb{S}^{n-1}$!

(i) Let $\mathbf{x}^* \in \mathbb{S}^{n-1}$ be such that $\lambda^* \neq 0$. Then letting $\mathbf{z} = (z_j)$,

$$\mathbf{x}^* = \sum_{j=1}^{m} z_j \, \boldsymbol{\ell}_j \quad \text{and} \quad \mathbf{y}^* := \mathbf{L}\,\mathbf{z} \Rightarrow \|\mathbf{y}^*\| = 1\,,$$

so that $\mathbf{y}^* \in \mathscr{E}_n$ and $f((\mathbf{L}_1 \cdot \mathbf{y}^*), \ldots, (\mathbf{L}_m \cdot \mathbf{y}^*)) = h(\mathbf{x}^*)$.

(ii) If $\lambda^* = 0$ then $\nabla f(\boldsymbol{\ell}\,\mathbf{x}^*) = 0$.

- Write $\mathbf{x}^* = \boldsymbol{\ell}^T\mathbf{u} \oplus \mathbf{v}$ with $\mathbf{v} \in \mathrm{Ker}(\boldsymbol{\ell})$, so that $\boldsymbol{\ell}^T\mathbf{u} \perp \mathbf{v}$.
- $1 = \|\mathbf{x}^*\| \Rightarrow \|\boldsymbol{\ell}^T\mathbf{u}\| \leq 1$, and $\mathbf{y}^* := \mathbf{L}\,\mathbf{u}$ implies $\|\mathbf{y}^*\| \leq 1$, so that $\mathbf{y}^* \in \mathscr{E}_m$.
- $\boldsymbol{\ell}\,\mathbf{x}^* = \boldsymbol{\ell}\,\boldsymbol{\ell}^T\mathbf{u} = \mathbf{L}\,\mathbf{L}\,\mathbf{u} = \mathbf{L}\,\mathbf{y}^*$, and therefore

$$h(\mathbf{x}^*) = f(\boldsymbol{\ell}\,\mathbf{x}^*) = f(\mathbf{L}\,\mathbf{y}^*)\,, \quad \text{with } \mathbf{y}^* \in \mathscr{E}_m.$$

For the converse, similar calculations yield the desired result.

Hence

solving **P** on $\mathbb{S}^{n-1}$ is equivalent to solving **Q** on $\mathscr{E}_m$ and the latter problem is lower-dimensional on the CONVEX SET $\mathscr{E}_m$, a significant progress.

The case $m = 1$

If $m = 1$ (so that $\ell \in \mathbb{R}^n$) then Problem **P** reduces to the easy to solve univariate problem

$$\mathbf{Q}: \quad f^* = \min_y \{ f(\|\ell\|\, y) : y \in [-1, 1] \}$$

Hence

solving **P** on $\mathbb{S}^{n-1}$ is equivalent to solving **Q** on $\mathscr{E}_m$ and the latter problem is lower-dimensional on the CONVEX SET $\mathscr{E}_m$, a significant progress.

The case $m = 1$

If $m = 1$ (so that $\boldsymbol{\ell} \in \mathbb{R}^n$) then Problem **P** reduces to the easy to solve univariate problem

$$\mathbf{Q}: \quad f^* = \min_y \{ f(\|\boldsymbol{\ell}\|\, y) : \, y \in [-1, 1] \}$$

The case $m = 1$ (continued)

If $m = 1$ and h is a quasi-convex polynomial then

$$f^* \;=\; \min\,[\,f(\|\boldsymbol{\ell}\|)\,,\,f(-\|\boldsymbol{\ell}\|)\,]$$

and so h has no spurious local minimum on $\mathbb{S}^{n-1}$.

☞ h quasi-convex implies $h(\mathbf{x}) = f(\boldsymbol{\ell}\cdot\mathbf{x})$ for some $\boldsymbol{\ell} \in \mathbb{R}^n$ and some monotonic univariate polynomial f (Ahmadi et al. (2013), Math. Program.).

Finally if h has the low-rank formulation

$$h(\mathbf{x}) = f(\boldsymbol{\ell}\,\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n,$$

for some *hidden* $\boldsymbol{\ell} \in \mathbb{R}^{m \times n}$ then by sufficiently many evaluations of ∇f at randomly chosen points $(\mathbf{x}(i))_{i \in J}$ (e.g. $\mathbf{x}(i) \in \mathbb{S}^{n-1}$) one may recover $\boldsymbol{\ell}$.

Happy Birthday !