The Geometry of Sparse Analysis Regularization

Samuel Vaiter
CNRS & IMB

Joint work with Xavier Dupuis (IMB)

December 4, 2019
@PGMO'19, Palaiseau

Inverse problem

Inverse problem / regression setting

$$y = \mathbf{\Phi} x_0 + w$$

observations

$\mathbb{R}^p$

noise

degradation operator

$\mathbb{R}^n \to \mathbb{R}^p$

ground truth

$\mathbb{R}^n$

Inverse problem

Inverse problem / regression setting

$$y = \mathbf{\Phi} x_0 + w$$

observations $\mathbb{R}^p$

noise

degradation operator $\mathbb{R}^n \to \mathbb{R}^p$

ground truth $\mathbb{R}^n$

Variational methods

compromise

$$x^\star \in \underset{x \in \mathbb{R}^n}{\mathrm{Argmin}} \; \boxed{F(\mathbf{\Phi} x, y)} + \lambda \boxed{J(x)}$$

data fidelity

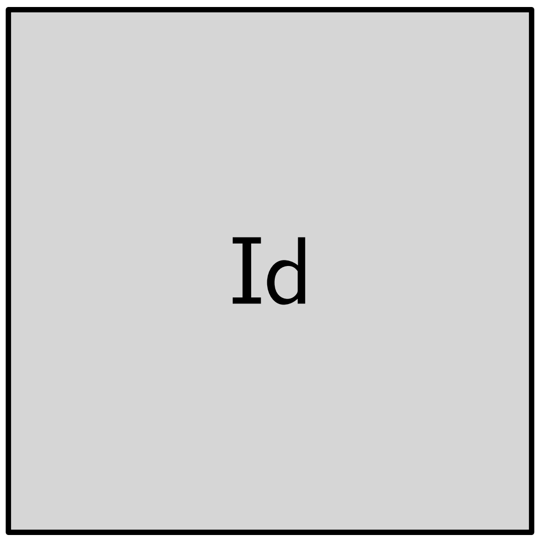
convex "regularization"

BPDN / Lasso
Total Variation
Nuclear
...

Dictionary / Analysis operator

Dictionary of $\mathbb{R}^n$: $\{d_i\}_{i=0}^{p-1}$

identity



shift invariant wavelet frame

finite difference operator

$$\begin{pmatrix} -1 & & & 0 \\ +1 & -1 & & \\ & +1 & \ddots & \\ & & \ddots & -1 \\ 0 & & & +1 \end{pmatrix}$$

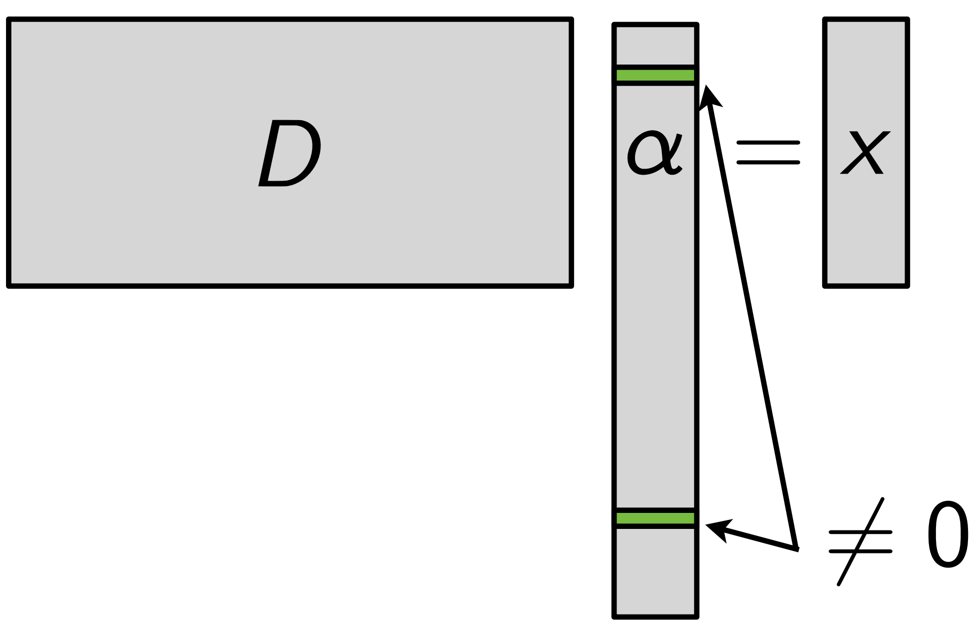
fused Lasso

Sparse regularizations

Synthesis

$$\operatorname*{Argmin}_{\alpha \in \mathbb{R}^q} \frac{1}{2}\|y - \Psi\alpha\|_2^2 + \lambda\|\alpha\|_1$$

$$\Psi = \Phi D \qquad x = D\alpha$$



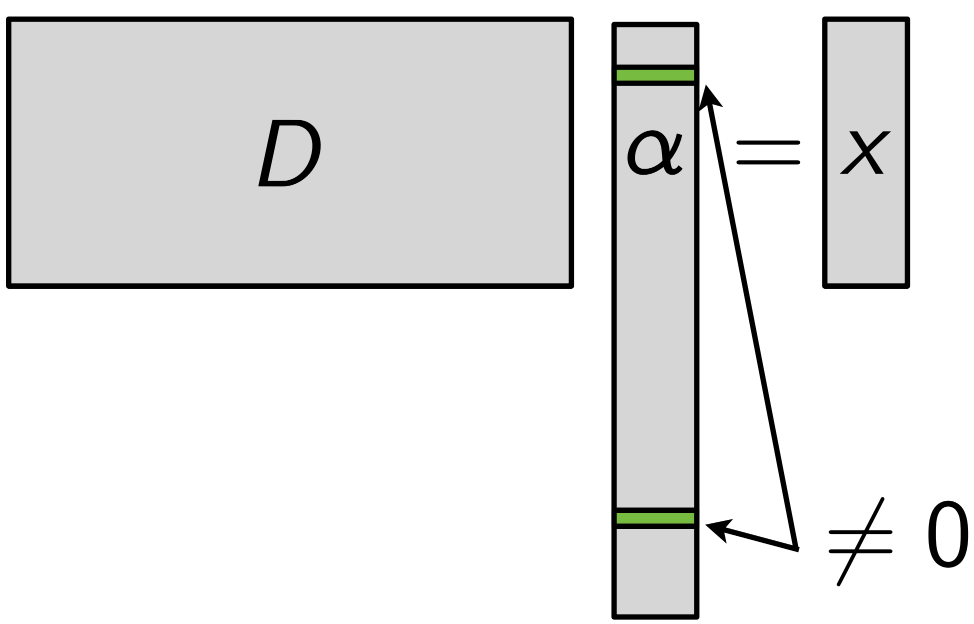
Sparse approx. of x in D

Sparse regularizations

Synthesis

$$\underset{\alpha\in\mathbb{R}^q}{\mathrm{Argmin}}\ \frac{1}{2}\|y-\Psi\alpha\|_2^2+\lambda\|\alpha\|_1$$

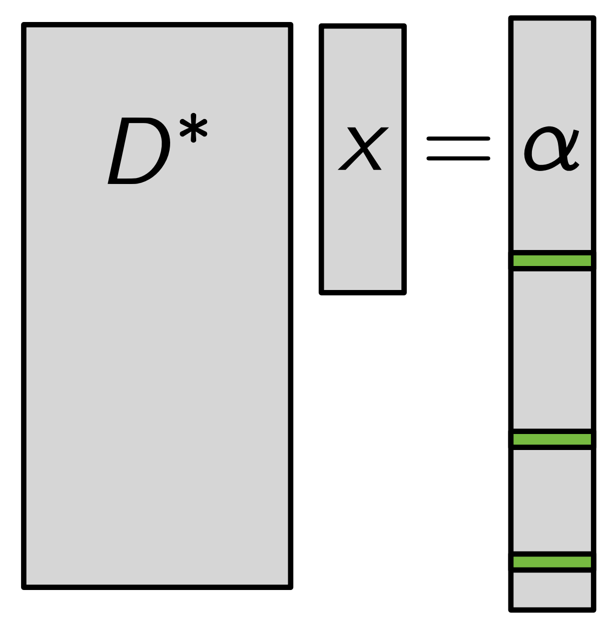
$$\Psi=\Phi D \qquad x=D\alpha$$



Sparse approx. of x in D

Analysis

$$\underset{x\in\mathbb{R}^n}{\mathrm{Argmin}}\ \frac{1}{2}\|y-\Phi x\|_2^2+\lambda\|D^* x\|_1$$



Correlation of x and D sparse

Uniqueness

$$\mathcal{S} = \underset{x \in \mathbb{R}^n}{\mathrm{Argmin}}\ \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

In a large number of situation,

$$\mathcal{S} = \{x^\star\}$$

[Nam et al. 2013, Vaiter et al. 2013, Zhang et al. 2016, Ali & Tibshirani 2018, …]

Uniqueness

$$\mathcal{S} = \operatorname*{Argmin}_{x \in \mathbb{R}^n} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

In a large number of situation,

$$\mathcal{S} = \{x^\star\}$$

[Nam et al. 2013, Vaiter et al. 2013, Zhang et al. 2016, Ali & Tibshirani 2018, …]

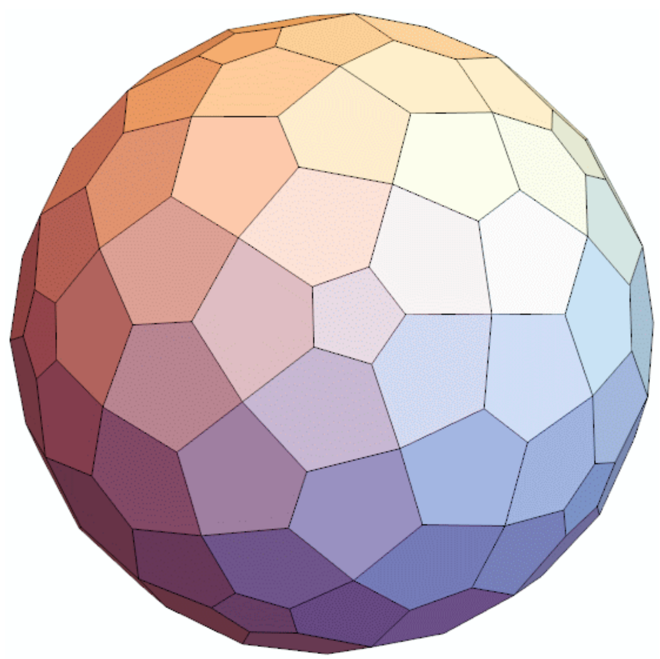
What about non-generic settings?

General results [Mangasarian 1988, Burke & Ferris 1991, …] on the geometry convex programs does not provide much

The solution set is a polytope

$$\mathcal{S} = \operatorname*{Argmin}_{x \in \mathbb{R}^n} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

Whatever Φ, D, y, the solution set $\mathcal{S}$ is a polytope



It may be unbounded (containing rays)

In theory, general results on polytopes e.g. [Zieglier 1995] may be applied

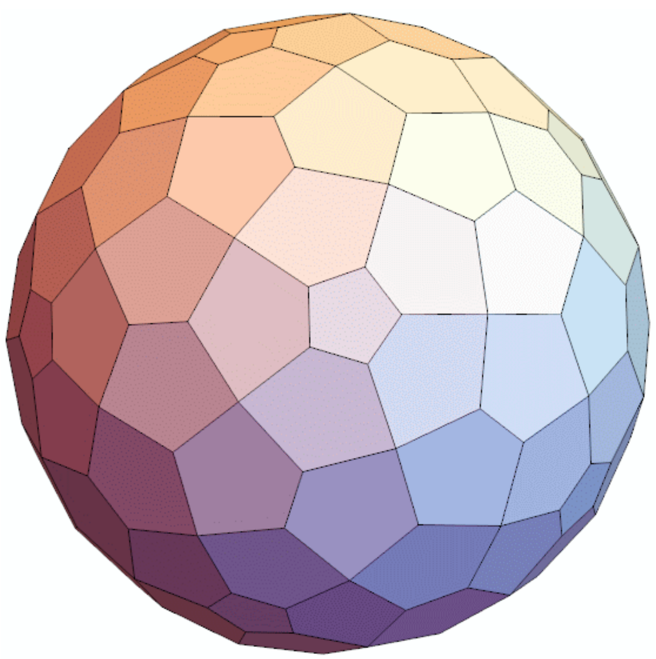
- description of faces
- signs matroids
- …

The solution set is a polytope

$$\mathcal{S} = \operatorname*{Argmin}_{x \in \mathbb{R}^n} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

Whatever Φ, D, y, the solution set $\mathcal{S}$ is a polytope

It may be unbounded (containing rays)



$\mathcal{S}$ is bounded

$$\Updownarrow$$

$$\operatorname{Ker} \Phi \cap \operatorname{Ker} D^* = \{0\}$$

Representer theorems

$$x^\star \in \operatorname*{Argmin}_{x \in \mathbb{R}^n} F(\mathbf{\Phi} x, y) + \lambda J(x)$$

$$\Updownarrow$$

$$x^\star = \sum_{i=1}^{n} \alpha_i \psi_i + u$$

extreme points
of the reg. level set

[Schölkopf et al. 2001, **Boyer et al. 2019**, Unser 2019, …]

Characterization of the solution set

$$\mathcal{S} = \underset{x \in \mathbb{R}^n}{\mathrm{Argmin}} \ \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

Solution set is an affine subspace intersected by the unit ball

For any Φ, y, λ, $\mathcal{S}$ is a nonempty polyhedron s.t. $\mathcal{S} = \mathcal{A} \cap B_r$

with $r \geqslant 0$ and $\mathcal{A}$ an aff. subspace such that $\emptyset \neq \mathcal{A} \cap B_r \subset \partial B_r$

Any affine subspace intersected by the unit ball is a sol. set

Let $r \geqslant 0$, $\mathcal{A}$ an aff. subspace such that $\emptyset \neq \mathcal{A} \cap B_r \subset \partial B_r$

There exists Φ, y, λ such that $\mathcal{S} = \mathcal{A} \cap B_r$ and $\mathrm{Ker}\,\Phi = dir(\mathcal{A})$

Support and signal model

$$\mathcal{S} = \underset{x \in \mathbb{R}^n}{\mathrm{Argmin}} \ \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

$$I = \mathrm{supp}(D^* x^\star), J = I^c$$

support

cosupport

Support and signal model

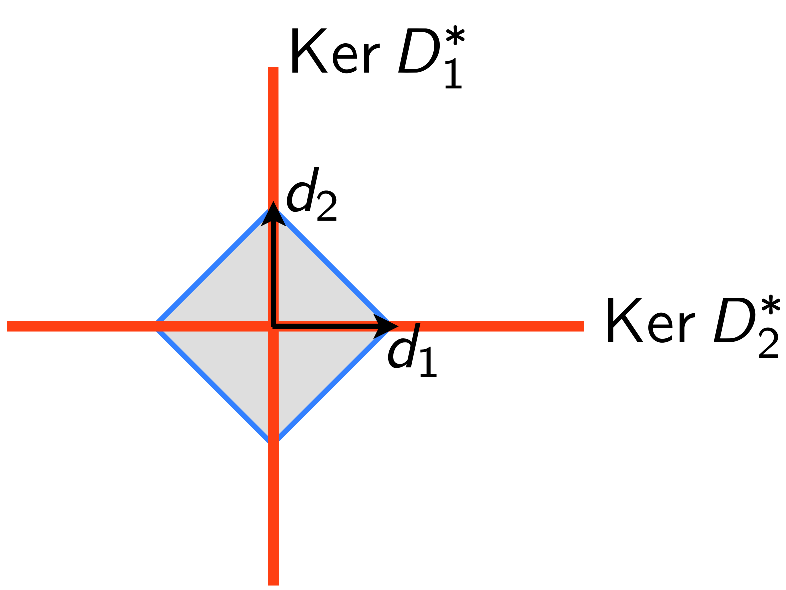
$$\mathcal{S} = \underset{x \in \mathbb{R}^n}{\mathrm{Argmin}} \ \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

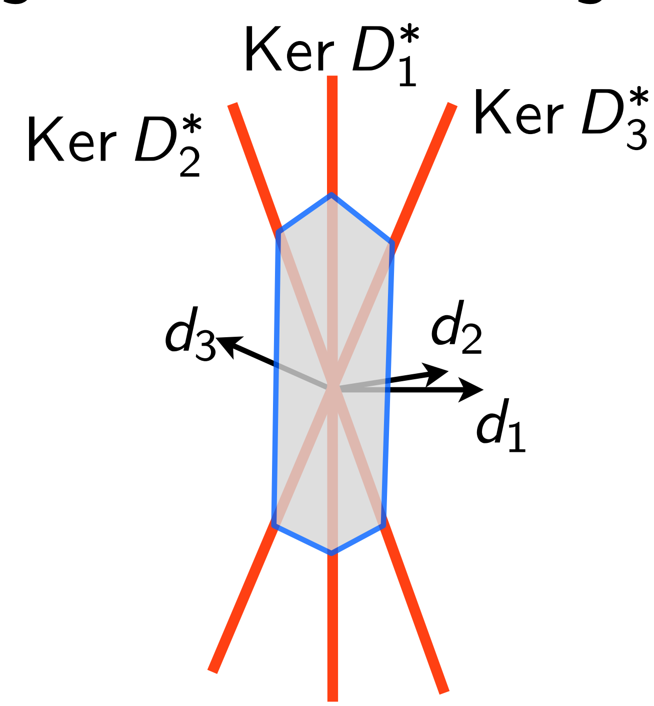
$$I = \mathrm{supp}(D^* x^\star), \ J = I^c$$

support — cosupport

Signal model: “union of subspace”

$$\Theta = \bigcup_{k \in 1 \dots q} \Theta_k \quad \text{where} \quad \Theta_k = \{\mathrm{Ker}\, D_J^* \ : \ \dim \mathrm{Ker}\, D_J^* = k\}$$





Extreme points of the solution set

$$\mathcal{S} = \underset{x \in \mathbb{R}^n}{\mathrm{Argmin}} \ \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

$$I = \mathrm{supp}(D^* x^\star), J = I^c$$

$x^\star \in \mathcal{S}$ is extremal iff $\mathrm{Ker}\, \Phi \cap \mathrm{Ker}\, D_J^* = \{0\}$

Extreme points of the solution set

$$\mathcal{S} = \underset{x \in \mathbb{R}^n}{\operatorname{Argmin}} \ \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

$$I = \operatorname{supp}(D^* x^\star), J = I^c$$

$x^\star \in \mathcal{S}$ is extremal iff $\operatorname{Ker}\Phi \cap \operatorname{Ker} D_J^* = \{0\}$

(H_J)-procedure

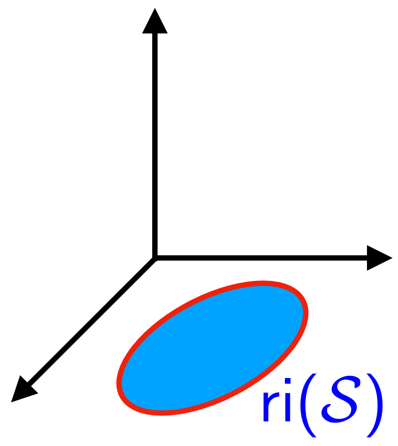
1. Start from a non-extremal point x
2. Find a direction in $z \in \operatorname{Ker}\Phi \cap \operatorname{Ker} D_J^*$
3. Find the smallest $x + tz$ with a different support
4. Repeat until (H_J) is satisfied

Interior and maximal solutions

$$\mathcal{S} = \underset{x \in \mathbb{R}^n}{\mathrm{Argmin}} \ \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

$$x^\star \in \mathrm{ri}(\mathcal{S}) \Longleftrightarrow x^\star \text{ is a maximal solution}$$

$\mathrm{ri}(\mathcal{S})$: relative interior of $\mathcal{S}$



maximal solution

$\forall x \in \mathcal{S}, \mathrm{supp}(D^* x^\star) \supseteq \mathrm{supp}(D^* x)$

Interior and maximal solutions

$$\mathcal{S} = \operatorname*{Argmin}_{x \in \mathbb{R}^n} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

$$x^\star \in \mathrm{ri}(\mathcal{S}) \iff x^\star \text{ is a maximal solution}$$

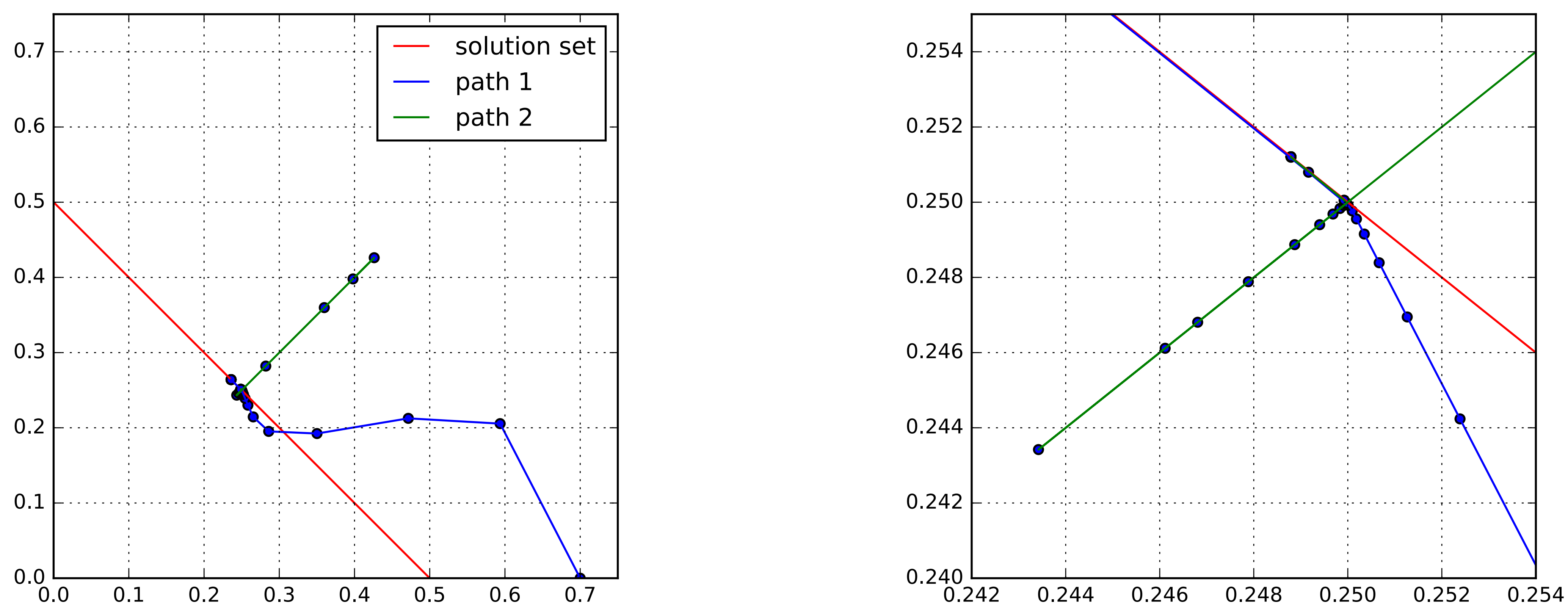
There exists an interior-point type algorithm to compute a maximal solution (not really scalable)

Open problem: condition on Chambolle—Pock (ADMM) to converge to a maximal solution

Interior and maximal solutions

$$\mathcal{S} = \underset{x \in \mathbb{R}^n}{\mathrm{Argmin}} \ \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

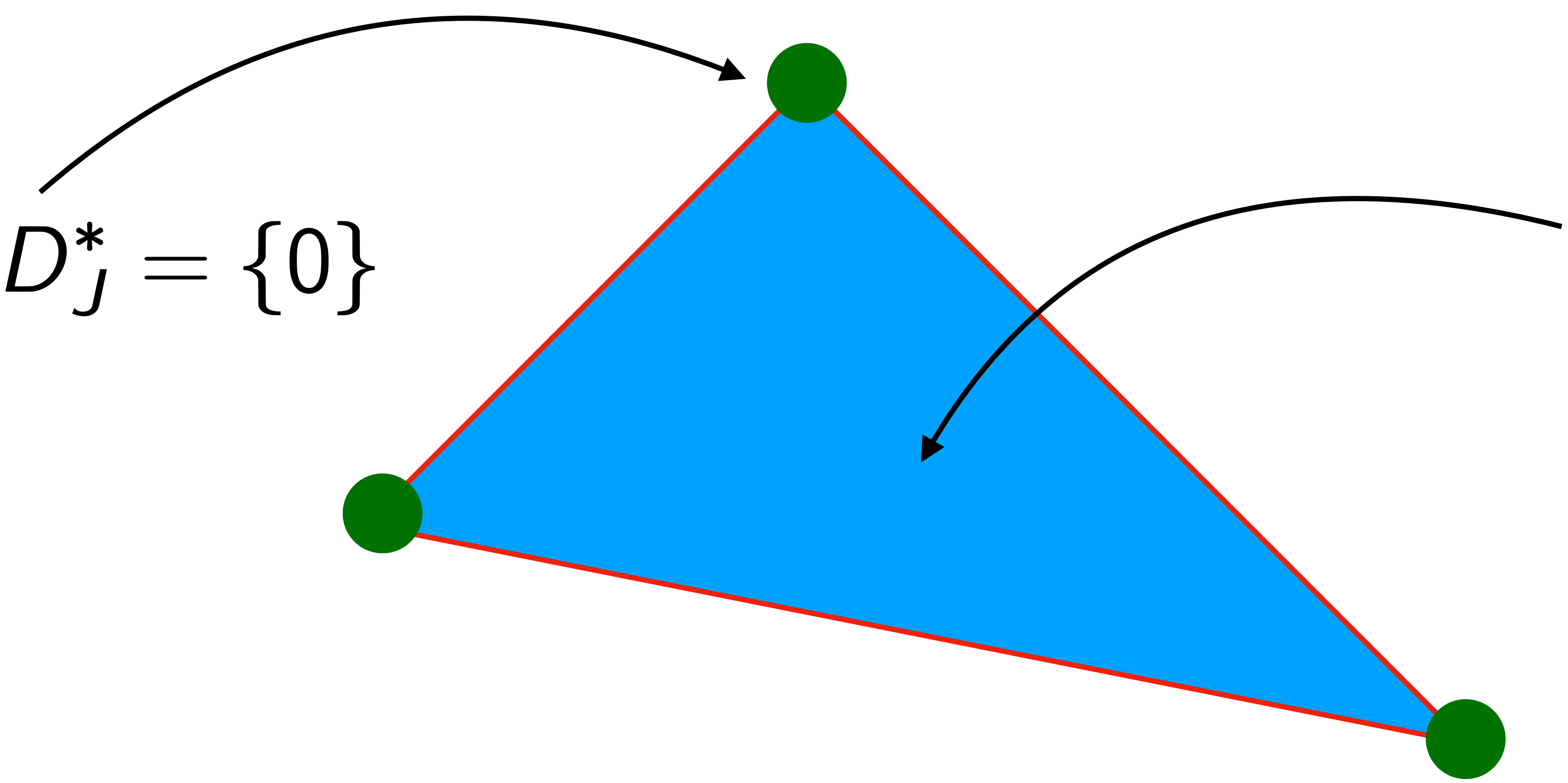
$x^\star \in \mathrm{ri}(\mathcal{S}) \iff x^\star$ is a maximal solution



Interior and extremal solutions

$$\mathcal{S} = \underset{x \in \mathbb{R}^n}{\mathrm{Argmin}} \ \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

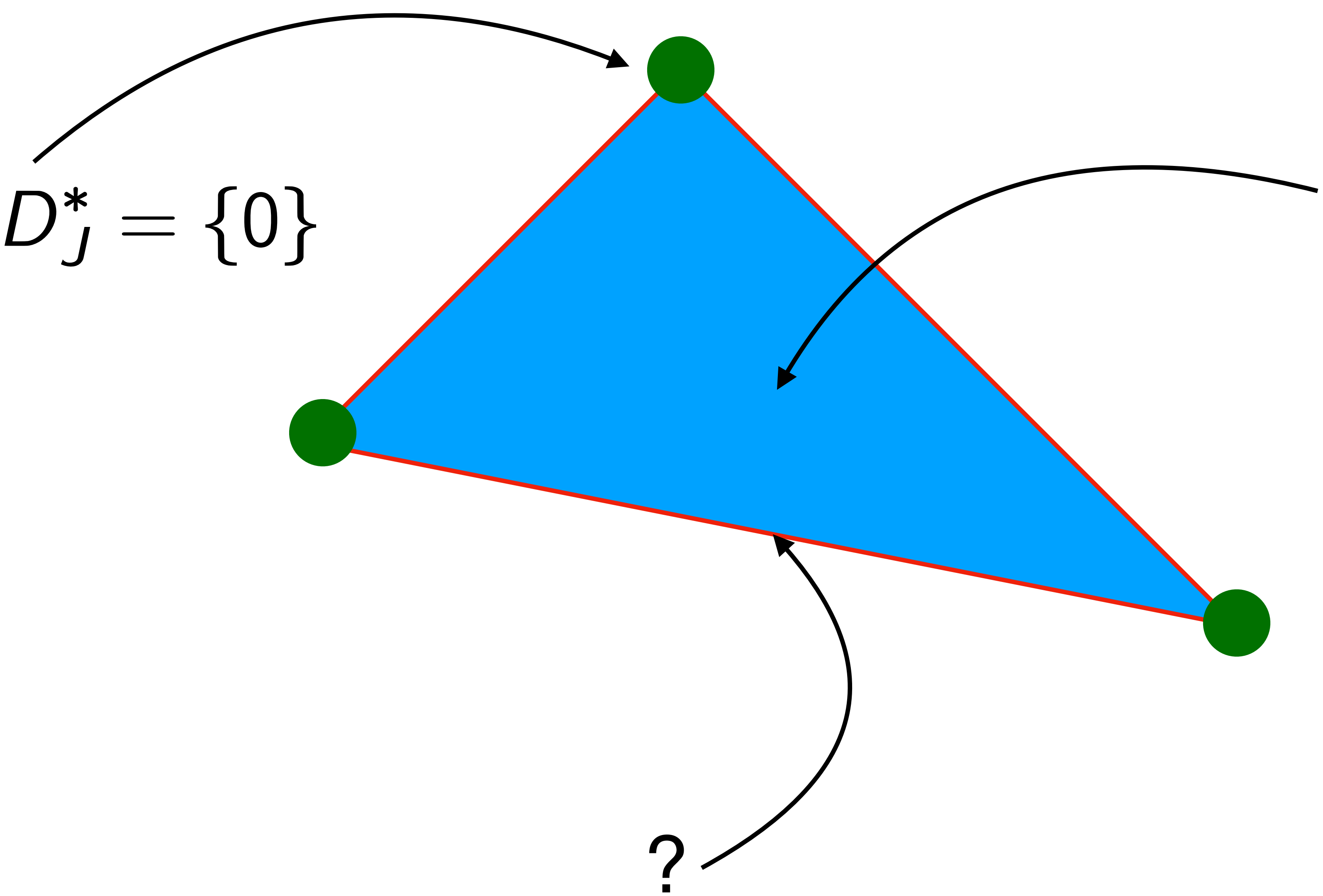
$$I = \mathrm{supp}(D^* x^\star), J = I^c$$



Interior and extremal solutions

$$\mathcal{S} = \underset{x \in \mathbb{R}^n}{\mathrm{Argmin}} \ \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

$$I = \mathrm{supp}(D^* x^\star), J = I^c$$



Description of the faces

$$\mathcal{S} = \underset{x \in \mathbb{R}^n}{\operatorname{Argmin}} \ \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

$\bar{x} \in \mathrm{ri}(X)$, $\bar{s} = \mathrm{sign}(D^* \bar{x})$, $r = \|D^* \bar{x}\|_1$, $\bar{F} = B_r \cap \{x : \langle D\bar{s},\ x\rangle = r\}$

Solution set description

$$X = (\bar{x} + \mathrm{Ker}\,\Phi) \cap \bar{F}$$

$$X = (\bar{x} + \mathrm{Ker}\,\Phi) \cap \{x \ : \ \mathrm{sign}(D^* x) \leqslant \bar{s}\}$$

$$\mathrm{ri}(X) = (\bar{x} + \mathrm{Ker}\,\Phi) \cap \{x \ : \ \mathrm{sign}(D^* x) = \bar{s}\}$$

$$\mathrm{dir}(X) = \mathrm{Ker}\,\Phi \cap \mathrm{Ker}\, D_J^* \text{ with } J = \mathrm{cosupp}(D^* \bar{x})$$

Faces description, any face of the solution set is of the form

$$F = X \cap \{x \ : \ J \subseteq \mathrm{cosupp}(D^* x)\}$$

$$\mathrm{ri}(F) = X \cap \{x \ : \ J = \mathrm{cosupp}(D^* x)\}$$

$$\mathrm{dir}(F) = \mathrm{Ker}\,\Phi \cap \mathrm{Ker}\, D_J^*$$

Toy use

$$\mathcal{S} = \underset{x \in \mathbb{R}^n}{\text{Argmin}} \ \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

Linear characterization: use of LP solver

$$\bar{x} \in \text{ri}(X), \bar{s} = \text{sign}(D^* \bar{x}), \bar{J} = \text{cosupp}(\bar{s}), \bar{I} = \text{supp}(\bar{s})$$

$$X = \left\{ x \ : \ \Phi x = \Phi \bar{x}, D^*_{\bar{J}} x = 0, \text{diag}(\bar{s}_{\bar{I}}) D^*_{\bar{I}} x \geqslant 0 \right\}$$

Toy use

$$\mathcal{S} = \underset{x\in\mathbb{R}^n}{\operatorname{Argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda\|D^* x\|_1$$

Linear characterization: use of LP solver

$$\bar{x} \in \operatorname{ri}(X), \bar{s} = \operatorname{sign}(D^*\bar{x}), \bar{J} = \operatorname{cosupp}(\bar{s}), \bar{I} = \operatorname{supp}(\bar{s})$$

$$X = \left\{x \ : \ \Phi x = \Phi\bar{x}, D^*_{\bar{J}} x = 0, \operatorname{diag}(\bar{s}_{\bar{I}}) D^*_{\bar{I}} x \geqslant 0\right\}$$

Find small/biggest coefficient

$$\max_{x\in\mathcal{S}} \langle x, \ e_i\rangle$$

Find small/biggest correlation

$$\max_{x\in\mathcal{S}} \langle x, \ d_i\rangle$$

obtain dispensable coefficient

Thanks for your attention!

- SV, G. Peyré, C. Dossal, J.M. Fadili. *Robust Sparse Analysis Regularization*. IEEE Trans. Inform. Theory, 2013.
- A. Barbara, A. Jourani, SV. *Maximal solutions of sparse analysis regularization*. J. Optim. Theory. Appl., 2019.
- **X. Dupuis, SV. *The Geometry of Sparse Analysis Regularization*. arXiv e-prints, 2019.**

PhD position available
in Dijon on Graphs,
Optimization and ML !