Extensions of Semidefinite Programming Methods for Atomic Decomposition

Hsiao-Han Chao   Lieven Vandenberghe

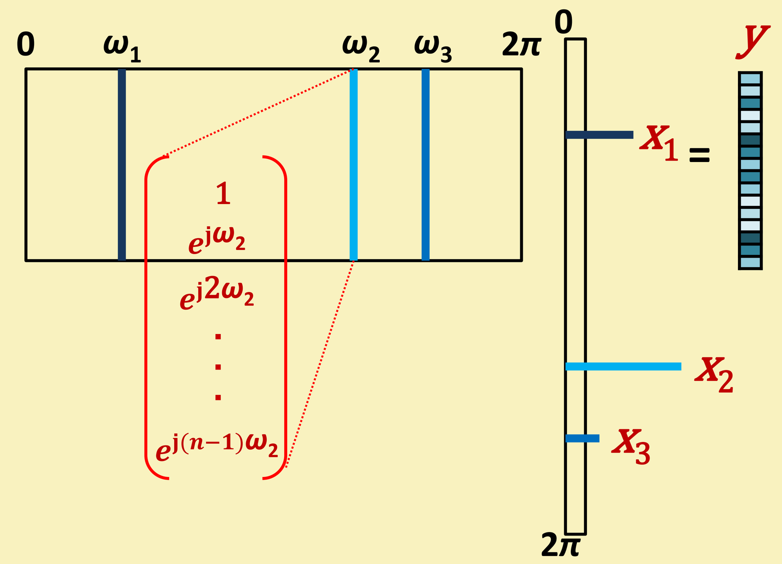
Department of Electrical Engineering, UCLA
suzihchao@ucla.edu, vandenbe@ucla.edu

1-norm Regularization

Minimize $f(y) + \sum_{k=1}^{r} |x_k|$

Subject to $y = \sum_{k=1}^{r} x_k a_k$

$a_k \in \mathcal{D},\ k = 1, \ldots, r$



- variables: y, $a_k \in \mathbb{C}^n$, $x_k \in \mathbb{C}$, r
- f : convex (penalty or indicator)
- $\mathcal{D}$: set of atoms (dictionary)
- extends lasso, (noisy) basis pursuit, etc. to non-finite $\mathcal{D}$

Matrix Pencil Dictionary

$$\mathcal{D} = \{a \in \mathbb{C}^n \mid \|a\|_2 = 1,\ (\lambda G - F)a = 0,\ \lambda \in \mathcal{C}\}$$

- $\lambda G - F$ is a $p \times n$ matrix pencil
- $\mathcal{C}$: a segment of a line or circle in the complex plane
- normalizations other than $\|a\|_2 = 1$ are possible
- example: complex exponentials

$$a = \frac{c}{\sqrt{n}}(1, e^{j\omega}, e^{j2\omega}, \ldots, e^{j(n-1)\omega}), \quad |c| = 1$$

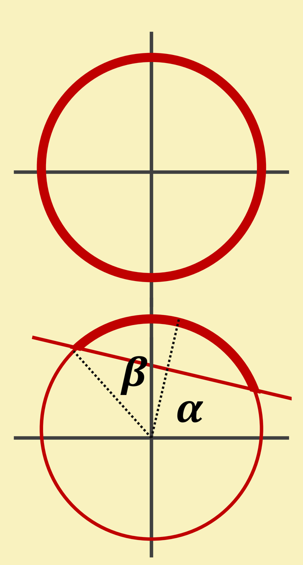
take $\mathcal{C}$ the unit circle, $F = \begin{bmatrix} 0 & I_{n-1} \end{bmatrix}$, and $G = \begin{bmatrix} I_{n-1} & 0 \end{bmatrix}$

Set $\mathcal{C}$

$$\mathcal{C} = \left\{ \lambda \in \mathbb{C} \cup \{\infty\} \mid \begin{bmatrix} \lambda \\ 1 \end{bmatrix}^H \Phi \begin{bmatrix} \lambda \\ 1 \end{bmatrix} = 0,\ \begin{bmatrix} \lambda \\ 1 \end{bmatrix}^H \Psi \begin{bmatrix} \lambda \\ 1 \end{bmatrix} \le 0 \right\}$$

- Φ and Ψ are 2×2 Hermitian matrices with $\det \Phi < 0$
- $\mathcal{C}$ includes $\lambda = \infty$ when $\Phi_{11} = 0$ and $\Psi_{11} \le 0$
- $\mathcal{C}$ is a circle or straight line if $\Psi = 0$
- examples:

Φ	Ψ	set $\mathcal{C}$
$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	0	$\lambda = e^{j\omega}$ (unit circle)
$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 0 & -e^{j\alpha} \\ -e^{-j\alpha} & 2\cos\beta \end{bmatrix}$	$\lambda = e^{j\omega},\ \lvert\omega - \alpha\rvert \le \beta$



Structured Low-rank Decomposition

Minimize $f(Y) + \sum_{k=1}^{r} \|x_k\|_2$

Subject to $Y = \sum_{k=1}^{r} a_k x_k^H, \quad a_k \in \mathcal{D},\ k = 1, \ldots, r$ (1)

- variables: $Y \in \mathbb{C}^{n\times m}$, $a_k \in \mathbb{C}^n$, $x_k \in \mathbb{C}^m$, r
- equivalent SDP with variables $Y \in \mathbb{C}^{n\times m}$, $X \in \mathbb{H}^n$, $Z \in \mathbb{H}^m$

Min. $f(Y) + (\mathbf{tr}\, X + \mathbf{tr}\, Z)/2$

S.t. $\begin{bmatrix} X & Y \\ Y^H & Z \end{bmatrix} \succeq 0$ (2)

$$\Phi_{11}FXF^H + \Phi_{21}FXG^H + \Phi_{12}GXF^H + \Phi_{22}GXG^H = 0$$
$$\Psi_{11}FXF^H + \Psi_{21}FXG^H + \Psi_{12}GXF^H + \Psi_{22}GXG^H \preceq 0$$

Lemma: Conic Decomposition

For $X \succeq 0$ with rank r satisfying the last two constraints in (2)

$$\Rightarrow \quad X = \sum_{k=1}^{r} \theta_k a_k a_k^H, \quad \theta_1, \ldots, \theta_r > 0, \quad a_1, \ldots, a_r \in \mathcal{D}$$

- follows from decompositions in the proofs of KYP lemma (PV13)
- efficiently computed via SVD and Schur decompositions

Proof of Equivalence

- SDP as relaxation of (1): if a_k, x_k, r are feasible in (1) then

$$\begin{bmatrix} X & Y \\ Y^H & Z \end{bmatrix} = \sum_{k=1}^{r} \|x_k\|_2 \begin{bmatrix} a_k \\ x_k/\|x_k\|_2 \end{bmatrix} \begin{bmatrix} a_k \\ x_k/\|x_k\|_2 \end{bmatrix}^H$$

is feasible in (2) with the same objective value

- exactness of relaxation: decompose feasible X of rank r as

$$X = \sum_{k=1}^{r} \theta_k a_k a_k^H, \quad \theta_k > 0, \quad a_k \in \mathcal{D}$$

first constraint in (2) implies there exist x_k such that

$$Y = \sum_{k=1}^{r} a_k x_k^H, \quad Z \succeq \sum_{k=1}^{r} \frac{1}{\theta_k} x_k x_k^H, \quad \frac{\mathbf{tr}\, X + \mathbf{tr}\, Z}{2} \ge \sum_{k=1}^{r} \|x_k\|_2$$

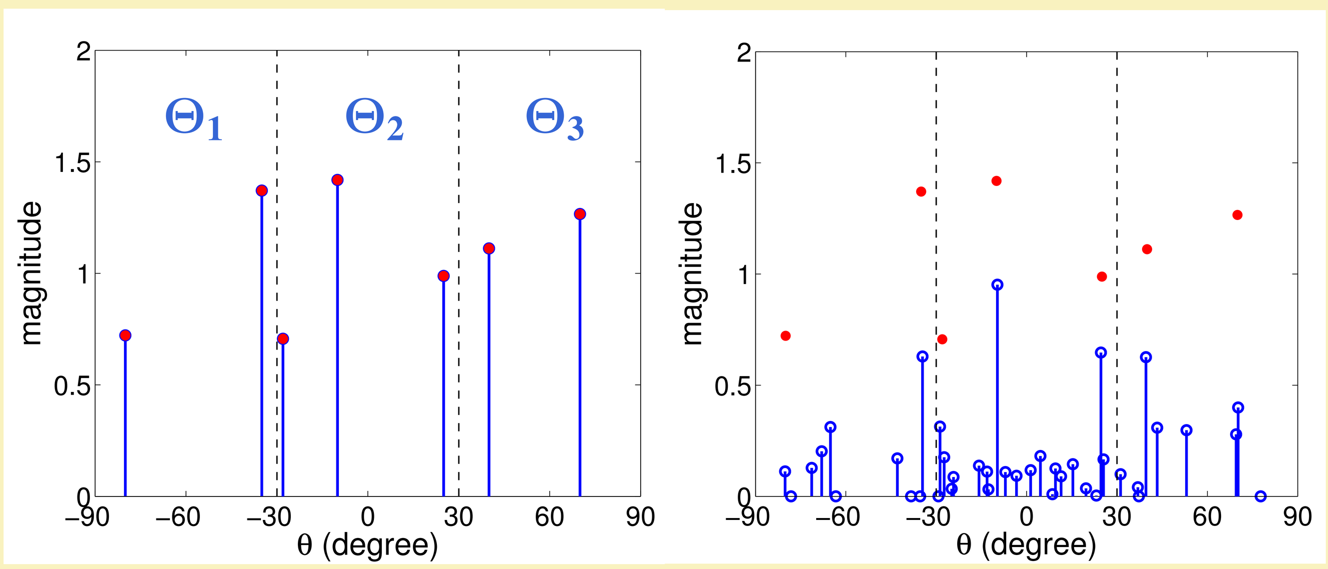
Example: DOA estimation

- sparse linear sensor array on a uniform grid of size $n = 50$
- use measurements b_1, b_2 from 2 groups of sensors
- group I_1 measures $\theta \in \Theta_1 \cup \Theta_2$, and group I_2, $\theta \in \Theta_2 \cup \Theta_3$
- can be solved via SDP formulation of:

Minimize $\sum_{j=1}^{3} \sum_{k=1}^{r_j} |x_{jk}|$

Subject to $y_j = \sum_{k=1}^{r_j} x_{jk} \begin{bmatrix} 1 \\ e^{j\omega(\theta_{jk})} \\ \vdots \\ e^{j(n-1)\omega(\theta_{jk})} \end{bmatrix}, \quad \omega(\theta) = \pi \sin\theta$

$$\theta_{jk} \in \Theta_j, \quad k = 1, \ldots, r_j,\ j = 1, 2, 3$$
$$(y_1 + y_2)_{I_1} = b_1, \quad (y_2 + y_3)_{I_2} = b_2$$



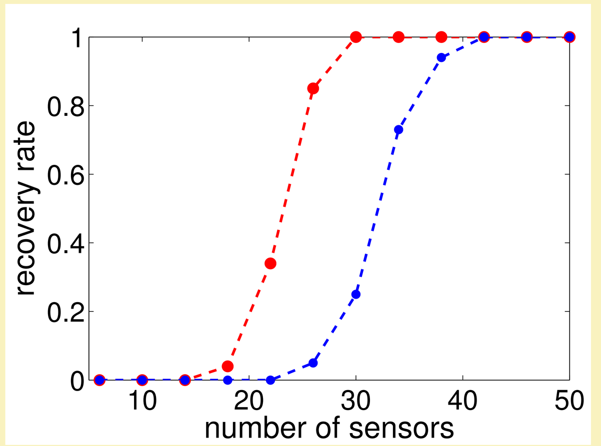
$|I_1| = |I_2| = 15$

red: true
blue: estimate

Figure : DOA estimation with and without interval constraints

- recovery rate against # of sensors $|I_1| + |I_2|$
- results averaged over 100 instances
- randomly generated magnitudes, I_1, I_2

red: with interval constraints
blue: no interval constraints



Key References

- E. J. Candès and C. Fernandez-Granda
  Towards a mathematical theory of super-resolution, 2014
- G. Tang, B. N. Bhaskar, P. Shah, and B. Recht
  Compressed sensing off the grid, 2013
- Z. Yang and L. Xie
  Exact joint sparse frequency recovery via optimization methods, 2014
- Y. Li and Y. Chi
  Off-the-grid line spectrum denoising and estimation with multiple measurement vectors, 2014
- K. V. Mishra, M. Cho, A. Kruger, and W. Xu
  Spectral super-resolution with prior knowledge, 2015
- T. Iwasaki and S. Hara
  Generalized KYP lemma: unified frequency domain inequalities with design applications, 2005
- G. Pipeleers and L. Vandenberghe
  Generalized KYP lemma with real data, 2011