Constructing an inexact dual certificate

Let $a_i \sim N(0, I_n)$.

Build $Y = \sum \lambda_i a_i a_i^t$ such that $Y_T \approx 0$ and $Y_{T^\perp} \lesssim -I_{T^\perp}$

Idea: $\frac{1}{m} \sum_{i=1}^m a_i a_i^t \approx \begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & & 1 \end{pmatrix}$

$\frac{1}{m} \sum_{i=1}^m a_i^2(1)\, a_i a_i^t \approx \begin{pmatrix} \beta & & 0 \\ & 1 & \\ 0 & & 1 \end{pmatrix}$ w/ $\beta = \mathbb{E}[z^4]$ where $z \sim N(0,1)$

So $\frac{1}{m} \sum_{i=1}^m (a_i^2(1) - \beta)\, a_i a_i^t \approx \begin{pmatrix} 0 & & & 0 \\ & 1-\beta & & \\ & & 1-\beta & \\ 0 & & & 1-\beta \end{pmatrix}$ ~~(get up to scale)~~

Try to build dual certificate as

$$Y = \frac{1}{m} \sum_{i=1}^m (a_i^2(1) - \beta)\, a_i a_i^t .$$

To show $Y_T \approx 0$ we need concentration estimate on $a_i^4(1)$.
Fourth power of gaussian is hard to control (not subexponential).
So, chop large values

$$Y = \frac{1}{m} \sum_{i=1}^m \left(a_i^2(1) \mathbb{1}_{|a_i(1)| \leq 3} - \beta_0\right) a_i a_i^t$$ w/ $\beta_0 = \mathbb{E}[z^4 \mathbb{1}_{|z| \leq 3}] \approx 2.67$

Now only second powers of Gaussians

Write $Y = Y^{(0)} - Y^{(1)}$

Concentration of Subexponentials

Lemma g (Bernstein)

Let X_i indep centeral subexponential rV w/ $K = \max_i \|X_i\|_{\psi_1}$

$$\forall\, \varepsilon \geq 0, \qquad P\left(\left| \frac{1}{N} \sum_{i=1}^{N} X_i \right| \geq \varepsilon \right) \leq 2e^{-c \min\left(\frac{\varepsilon^2}{K^2}, \frac{\varepsilon}{K} \right) N}$$

Thm: Let A be $N \times n$ w/ rows A_i that are indep subgaussian isotropic in $\mathbb{R}^n$. $\forall\, t \geq 0$ then w prob $1 - 2e^{-ct^2}$

$$\sqrt{N} - C\sqrt{n} - t \leq \sigma_{min}(A) \leq \sigma_{max}(A) \leq \sqrt{N} + C\sqrt{n} + t$$

where C depends on $K = \max_i \|A_i\|_{\psi_1}$

Recall: isotropic means $\mathbb{E}\, XX^t = I$

Lemma (Candès + Li)

Fix $x_0 = e_1$. With prob at least $1 - O(e^{-\gamma m})$, $\| Y_{T^\perp} + \frac{17}{10} I_{T^\perp} \| \leq \frac{1}{10}$

& $\| Y_T \|_F \leq \frac{3}{20}$.

Proof:

a) $\| Y^{(1)}_{T^\perp} - \beta_0 I_{T^\perp} \| \leq \frac{\beta_0}{40}$ w/ prob $1 - 2e^{-\gamma m}$ if $m \geq cn$

b) $\| Y^{(0)}_{T^\perp} - \alpha_0 I_{T^\perp} \| \leq \frac{\alpha_0}{40}$ when $\alpha_0 = \mathbb{E}[z^2 1_{|z| \leq 3}] \approx 0.97$
w/ prob $1 - 2e^{-\gamma m}$ if $m \geq cn$

c) $| Y(1,1) |^2 \leq \frac{1}{20}$ with prob. $1 - 2e^{-\gamma m}$ if $m \geq cn$

d) $\| Y_{(2:n,\,1)} \|_2^2 \leq \frac{1}{10}$ " " " " " "

a) Concentration of singular values of Gaussian matrices

b) Concentration of singular values of subGaussian matrices

$$Y^{(0)}_{T^\perp} = \frac{1}{m} \sum_{i=1}^m y_i y_i^t \quad \text{w/ } y_i := a_i(1)\, 1_{|a_i(1)| \leq 3}\, a_i' \quad \text{w/ } a_i = \begin{pmatrix} a_i(1) \\ | \\ a_i' \\ | \end{pmatrix}$$

So $\frac{y_i}{\sqrt{\alpha_0}}$ is isotropic

Can control Eigen values of $Y^0_{T^\perp}$ by sing values of $\begin{pmatrix} - y_1 - \\ - y_2 - \\ \vdots \\ - y_m - \end{pmatrix}$

c) $Y(1,1) = \frac{1}{m} \sum_{i=1}^m y_i$ where $y_i = \underbrace{\underbrace{1_{|z| \leq 3} z^4}_{\text{bounded}} - \underbrace{\beta_0 z^2}_{\text{squared Gaussian}}}_{\text{subGaussian}}$ for $z \sim N(0,1)$

Bernstein inequality

$$P\left(|Y(1,1)| \geq \frac{1}{\sqrt{20}} \right) \leq 2e^{-\gamma m} \quad \text{if } m \geq cn$$

d)

d) Let $y' = Y(2{:}n, 1)$ $\qquad Y = \begin{pmatrix} Y(1,1) & - y' - \\ | & \\ y' & Y_{r^\perp} \\ | & \end{pmatrix}$

$$y' = \frac{1}{m} \underbrace{\begin{pmatrix} - a_1' - \\ - a_2' - \\ \vdots \\ - a_m' - \end{pmatrix}}_{Z'} c \qquad \text{w/ } c_i := a_i(1)^3 \, \mathbb{1}_{|\langle a_i, x_0\rangle| \leq 3} - \beta_0 \, a_i(1)$$

Show c_i^2 are subexponential and concentrate by Bernstein

For fixed x of norm 1, $\|Z' x\|_2^2$ is χ^2 random var w/ $n-1$ d.o.f. and concentrates.

Numerical Solutions to PhaseLift

$$\min \ tr(X) \quad st \quad \mathcal{A}X = b, \quad X \succeq 0$$

can be written as a three-term optimization

$$\min_X \ \underbrace{\lambda tr(X)}_{\text{free parameter}} + \underbrace{\tfrac{1}{2}\|\mathcal{A}(X) - b\|^2}_{\text{soft data penalty}} + \underbrace{i_{X \succeq 0}(X)}_{\text{indicator function}}$$

can be written as two-term optimization

$$\min_X \ \underbrace{\lambda tr(X) + \tfrac{1}{2}\|\mathcal{A}(X) - b\|^2}_{F \text{ smooth}} + \underbrace{i_{X \succeq 0}(X)}_{G \text{ not smooth}}$$

Suggests projected gradient descent

$$X_{n+1} = P_{X \succeq 0}\left(X_n - \alpha \nabla F(X_n) \right)$$

Can be accelerated by Nesterov method.

Drawback: free parameter

Numerical Solutn to PhaseLift feasibility

Find X st $X \succeq 0$, $AX = b$.

$$\min_{X} \; i_{X \succeq 0}(X) + i_{A(X)=b}(X)$$

Solve by projection on convex sets

$$X_{n+1} = P_{X \succeq 0} P_{A(X)=b} X_n$$

where $\cancel{P_{X\succeq 0}(X) = \sum_{i=1}^{n}}$

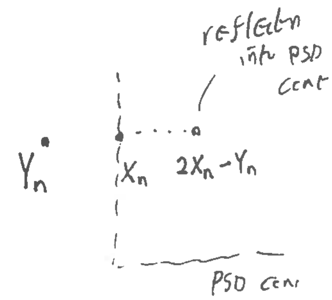
$$P_{X \succeq 0} : X = \sum_{i=1}^{n} \lambda_i v_i v_i^t \longmapsto \sum_{i=1}^{n} \max(\lambda_i, 0)\, v_i v_i^t$$

$$P_{AX=b} : X \longmapsto X - A^*(AA^*)^{-1}(AX - b)$$

Can be accelerated by Douglas-Rachford

$$Y_n = P_{AX=b}(\underbrace{2X_{n-1} - Y_{n-1}}_{\text{reflectn into PSD cone}}) - X_{n-1} + Y_{n-1}$$

$$X_n = P_{X \succeq 0} Y_n$$



Noisy Case

Let $b_i = a_i^t x_0 x_0^t a_i + \omega_i$

Find X_0 by solving

$$\min_{X \in \mathbb{R}^{n \times n}} \sum_{i=1}^m |a_i^t X a_i - b_i| \quad \text{st} \quad X \succeq 0 \qquad (*)$$

To get estimate for $x_0 \in \mathbb{R}^n$, take lead eigenvector of optimizer of $(*)$.

Thm: Let $a_i \sim N(0, I_n)$. $i = 1 \ldots m$

If $m \geq cn$, then w/ probability at least $1 - e^{-\gamma m}$

$\forall x_0$, the minimizer $\hat{X}$ to $(*)$ satisfies

$$\|\hat{X} - x_0 x_0^t\|_F \leq C_0 \frac{\|w\|_1}{m}$$

The resulting estimate ~~for~~ $\hat{x}_0$ satisfies

$$\|\hat{x} - e^{i\phi} x_0\|_2 \leq C_0 \frac{\|w\|_1}{m\|x_0\|_2} \quad \text{for some } \phi.$$

Recovery is robust to noise

Other comments: ~~E~~ Thm holds for $x_0 \in \mathbb{C}^n$ and

$$a_i \sim N(0, \tfrac{I_n}{2}) + i N(0, \tfrac{I_n}{2})$$

Phase Retrieval for sparse signals

Let $X_0 \in \mathbb{R}^n$ $\quad \|X_0\|_0 = k << n$

Let $a_i \sim N(0, I_n)$ $\quad i = 1 \dots m$

Let $b = |\langle a_i, X_0 \rangle|^2$.

Given a_i, b, find X_0 by a tractable algorithm

~~Use~~ ~~Try to~~ get

Open question: Is it possible to recover w/ $O(k)$ meas?

If we solve via phaselift

$$\min \operatorname{tr}(X) \quad \text{st} \quad \begin{matrix} X \succeq 0 \\ AX = b \end{matrix},$$ require $O(n)$ measurements

Try to penalize sparsity

$$\min \operatorname{tr}(X) + \lambda \|X\|_1 \quad \text{st} \quad \begin{matrix} X \succeq 0 \\ AX = b \end{matrix}$$ w/ $\|X\|_1$ is l_1 norm of vectorization of X.

Necessary ~~and suff~~ to have $\min(O(k^2), O(n))$ measurements.

In general, this problem is open.

Can solve in special cases

Eg if you have same relative phase of $\langle a_i, X_0 \rangle$ & $\langle a_j, X_0 \rangle$

can solve for overall phase. Solve CS problem.