Compressed sensing with convex optimization

Recall

Let $A \in \mathbb{R}^{m \times n}$ have independent entries satisfying:

$E A_{ij} = 0$, $E A_{ij}^2 = 1$, $\|A_{ij}\|_2 < 10$.

Let $0 \in K \subset \mathbb{R}^n$. Then

$E \sup_{x \in K} \|Ax\|_2 - \sqrt{m}\|x\|_2 \leq C \cdot w(k)$

$\Rightarrow P(\sup_{x \in K} \|Ax\|_2 - \sqrt{m}\|x\|_2 \leq C \cdot w(k)) \geq 99\%$,

Application: Signal recovery in structured linear model:

$y = Ax$, $x \in K \subset \mathbb{R}^n$

Recover or estimate $x$, by solving program:

Find $x' \in K$ s.t. $y = Ax'$

Let $\hat{x}$ be the solution.

Accuracy of $\hat{x}$?

Note: $\hat{x}, x \in K$, $A \hat{x} = Ax = y$.

Let $h = \hat{x} - x$. Then $h \in K - K$, $h \in \text{null}(A)$

$\|h\|_2$ is controlled via the low $m^*$ estimate, which follows from the above.

Thm (Low $m^*$ estimate)

Under the conditions of theorem above

$diam(\text{null}(A) \cap K) \leq C \cdot w(k)$, w.p. $\geq 99\%$
Thus, w.p. > 99%:

\[ \|h\|_2 \leq C \frac{w(k-k)}{\sqrt{m}} \]

Ex) \( k = \beta \),

\[ w(k-k) = w(2\beta) = E \sup_{x \in 2\beta} \langle x, g \rangle = E \sup_{x \in \beta} \langle x, g \rangle \approx C \frac{1}{\sqrt{m}} \]

\[ \|h\|_2 \leq C \frac{\log m}{\sqrt{m}} \]

\[ 2\beta = \]

\( \beta \)

Signal processing interpretation:

As soon as \( m \gg \sqrt{\log m} \), soln becomes very accurate.

Geometric interpretation: Intersection of \( B_1 \) w/ a random hyperplane of codimension \( m \gg \sqrt{\log m} \) has much smaller Euclidean diameter than that of \( B_1 \).

However, the approach above can be considerably improved.

Q: What if \( k \) is unbounded \( \Rightarrow w(k) = \infty ? \)

Q: What if the program (1) is computationally tractable?

Both problems occur when \( K_S = \{ x \in \mathbb{R}^n : \| x \|_0 \leq S \} \).
Solution: convexify and examine local properties of feasible set.

Motivating example:

\[ y = Ax, \quad x \in K \]

Let \( x \) be a soln to the convex program:

Find \( x' \) s.t. \( Ax = Ax' \), \( \|x'\| \leq \|x\| \),

Set \( h = x - x' \). Note:

1. \( h \in N(A) \)
2. \( \|x + h\| \leq \|x\| \),

Locally, the second constraint looks like a cone.

Define (tangent cone):

\[ D(K, x) := \{ v(x - x) : v \in K, v \geq 0 \} \]

Observe:

a) \( h \in D(K, x) \)
b) If \( D(K, x) \cap N(A) = \{ 0 \} \) then \( h = 0 \).
i.e. \(N(A)\) "escapes" the tangent cone.

The prob. of this good event is captured by the following theorem:

**Thm.** (Gordon's "escape through the mesh" thm.)

Let \(A\) be sub-Gaussian as in the "master theorem". Let \(D\) be a cone. Let \(c = c \frac{\text{w}(D \cap B^n)}{\sqrt{m}}\).

Then w.p. \(\geq 99\%\), \(\forall x \in D\)

A) \((1 - \varepsilon)\|x\|_2 \leq \frac{\|Ax\|_2}{\sqrt{m}} \leq (1 + \varepsilon)\|x\|_2\).

In particular, if \(m \geq \text{c} \text{w}(D \cap B^n)\), then

B) \(N(A) \cap D = \{0\}\) w.p. \(\geq 99\%\).

A) "\(A\) is well conditioned when restricted to \(D\) if \(m \geq \text{c} \text{w}(D \cap B^n)\)."

**Proof.** B) follows from the lower bound of A).

Exercise: Prove A) based on "master thm."

Return to example w/ \(x \in k^3, h \in D(0, 1, B^n, x)\).

One can show that

\[\text{w}(D(k, 1, B^n, x) \cap B^n) \leq C_5 \log \left(\frac{C_5}{5}\right)\]

Thus, "escape through mesh" thm implies that \(x = x\) w.p. \(\geq 99\%\) if \(m \geq C_5 \log (C_5)\).

Astonished researcher (2004): "\(x\) is recovered exactly! By a convex program! w/ hardly more measurements than would be needed for \(\ell_0\) minimization!"
General analysis using tangent cone:

Let $x$ be a sale to convex set containing $x$.

Find $x'$ s.t. $Ax' = y$, $x' \in K$ (2)

**Corollary** If $n \geq \omega(D(K,x) \cap B_{\varepsilon})^2$, then w.p. > 99%,$x' = x$.

**Proof:** Same steps as in example above.

**Remarks**

- (2) is often computationally tractable.

- $K$ is not necessarily the signal structure, rather it is a convex surrogate. Given a certain signal structure, determining what $K$ to use is a question of interest.

- In some cases $w(k,x)^2$ can be tiny. This is what allows mean, i.e., "dimension reduction".