3. Compressed Sensing

- We assume that an unknown object $x \in \mathbb{R}^N$ is measured by linear measurements, i.e., we can work with $\langle a_1, x\rangle, \dots, \langle a_m, x\rangle$ for some $a_1, \dots, a_m \in \mathbb{R}^N$.

In matrix notation $A = \begin{pmatrix} — a_1 — \\ \vdots \\ — a_m — \end{pmatrix}$ and we have $A \in \mathbb{R}^{m \times N}$ and $Ax \in \mathbb{R}^m$ available.

- N might be huge but its internal information does not grow...
  we assume, that $x \in \mathbb{R}^N$ is sparse, i.e. that $\|x\|_0 = \#\{j \in \{1, \dots, N\} : x_j \neq 0\}$ is small ... $\|x\|_0 \ll N$.

- The task of compressed sensing is to
  - design $A \in \mathbb{R}^{m \times N}$ and an algorithm $\Delta : \mathbb{R}^m \to \mathbb{R}^N$, such that for every $x \in \mathbb{R}^N$ with $\|x\|_0 \leq s$ $\Delta(Ax)$ is very close (or equal) to x. We want $m = m(s, N)$ as small as possible.

- The "traditional" sensing assumes that
  - $m = N$, $a_i = e_i$ = i-th canonical vector ... $A = I$, $\Delta = I$ & s plays no role

- We want $m \ll N$, actually $m \sim s$?!

Before we come to the modern part, we review the old stuff.

- Discrete Fourier transform

  Let $F = (F(0), \dots, F(N-1)) \in \mathbb{C}^N$ be given, then

  $\hat{F} = (\hat{F}(0), \dots, \hat{F}(N-1)) \in \mathbb{C}^N$ given by

  $$\hat{F}(k) = \sum_{n=0}^{N-1} F(n) \cdot \underbrace{\exp\left(-\frac{2\pi i}{N} kn\right)}_{\cos\left(\frac{2\pi}{N} kn\right) - i \sin\left(\frac{2\pi}{N} kn\right)}$$ is the DFT of F.

- Sometimes $\hat{F}$ is identified with $\{\hat{F}(k)\}_{k \in \mathbb{Z}}$, which is N-periodic.

Theorem (Fast Fourier Transform): Let $w_N \in \mathbb{C}$ with $w = \exp\left(-\frac{2\pi i}{N}\right)$ and $N = 2^m$ are given. Then $4 \cdot 2^m \cdot m = 4N \log_2(N) = O(N \log N)$ operations are enough to calculate $\hat{F}$.

Proof: Let $\#(M)$ denote the minimal number of operations needed to calculate the Fourier transform of a $F \in \mathbb{C}^M$. Then

(Claim) $\#(2M) \leq 2\#(M) + 8M$ if $w_{2M} = \exp\left(-\frac{2\pi i}{2M}\right)$ is given.

Then the Theorem follows by induction:

- $m = 1$, $N = 2^m$: $\#(2) = 4 \leq 8$ ... $\hat{F}(0) = F(0) + F(1)$; $\hat{F}(1) = F(0) - F(1)$
- $m-1 \to m$: $N = 2^{m-1}$: $\#(2N) \leq 2 \cdot \#(N) + 8N \leq 2 \cdot 4 \cdot 2^{m-1}(m-1) + 8 \cdot 2^{m-1}$
  $= 8m \cdot 2^{m-1} = 4m2^m$.

Proof of the Claim: • $2M$ operations for $w_{2M}^2, w_{2M}^3, \dots, w_{2M}^{2M-1}$

- $F_0(r) := F(2r)$, $F_1(r) = F(2r+1)$ ... even and odd part of F

  $0 \leq r \leq M-1$

Then $\hat{F}(k)=\sum_{r=0}^{2M-1}F(r)\,\omega_{2M}^{kr}=\sum_{l=0}^{M-1}F(2l)\,\omega_{2M}^{k(2l)}+\sum_{m=0}^{M-1}F(2m+1)\,\omega_{2M}^{k(2m+1)}$

$0\le k\le 2M-1$

$$=\sum_{l=0}^{M-1}F_0(l)\,\omega_M^{kl}+\omega_{2M}^{k}\sum_{m=0}^{M-1}F_1(m)\underbrace{\omega_{2M}^{2km}}_{\omega_M^{km}}$$

$$=\hat{F}_0(k)+\omega_{2M}^{k}\hat{F}_1(k)$$

... here $\hat{F}_0$ is periodic
$\hat{F}_0(k)=\hat{F}_0(k-M)$ for $k>M$.

Prony's method for sparse recovery

Let $F_N=\left(\exp\left(-\frac{2\pi i}{N}km\right)\right)_{k,m=0}^{N-1}$ be the Fourier matrix, i.e. $\hat{F}=F_N\cdot F$
$\hat{x}=F_N x$.

Theorem: Let $0\le s<N/2$. Put $m=2s$ and $A=$ "first $2s=m$ lines of F_N", $\in\mathbb{C}^{m\times N}$.

Then there is fast (= polynomial) algorithm, which recovers every s-sparse $x\in\mathbb{C}^N$ from $y=Ax$.

Proof: • First, we describe the algorithm

• Input: $s, N, y\in\mathbb{C}^{2s}$, $y(j)=\hat{x}(j)=\sum_{k=0}^{N-1}x(k)\exp\left(-2\pi i\frac{jk}{N}\right), j=0,\dots,2s-1$

• Solve the system

$$\begin{pmatrix} y(s-1), & y(s-2), & \dots, & y(0)\\ y(s), & y(s-1), & \dots, & y(1)\\ \vdots & & \ddots & \vdots\\ y(2s-2), & \dots & \dots, & y(s-1)\end{pmatrix}\begin{pmatrix} p(1)\\ p(2)\\ \vdots\\ p(s)\end{pmatrix}=-\begin{pmatrix} y(s)\\ y(s+1)\\ \vdots\\ y(2s-1)\end{pmatrix} \qquad (1)$$

and obtain $p(1),\dots,p(s)$. Put $p(0)=1$, $p(k)=0$ for $k>s$
$\to p\in\mathbb{C}^N$

• Put $S := (\operatorname{supp} z^\vee)^c$ ... the complement of the support of $z^\vee$
(the inverse Fourier transform of z)

We show that $\# S \le s$ and $\operatorname{supp} x \subset S$.

• Then just solve (2): $y(j) = \sum_{k \in S} x(k) \exp(-\frac{2\pi i}{N} jk)$, $j = 0, 1, \dots, 2s-1$ for $x(k)$.

Remarks: • (1) ... $s \times s$, linear system, $O(s^3)$

• find $\operatorname{supp} z^\vee$ ... $O(N \log N)$

• (2) ... $s \times 2s$ system ... $O(s^3)$ ... $\Rightarrow O(s^3 + N \log N)$

unstable, non-robust

Proof of the correctness of the algorithm

• Put $\tilde{S} = \operatorname{supp} x$, $\# \tilde{S} \le s$ ... assumption on sparsity

We want to show that $S = (\operatorname{supp} z^\vee)^c \supset \tilde{S} = \operatorname{supp} x$ for

every z with (1), $z(0) = 1$, $z(k) = 0$ for $k > s$.

Step 1: we show that (1) has at least one solution, and we

show the theorem if this solution is unique.

$$p(t) := \frac{1}{N} \prod_{k \in \tilde{S}} \left(1 - e^{-2\pi i \frac{k}{N}} \cdot e^{2\pi i \frac{t}{N}}\right), \quad t \in \{0, 1, \dots, N-1\}$$

$$p(t) = 0 \iff t \in \tilde{S} \quad \dots \quad \operatorname{supp} p = \{0, 1, \dots, N-1\} \setminus \tilde{S}.$$

Hence $p \cdot x = (p(k) \cdot x(k))_{k=0}^{N-1} \equiv 0$ ... hence $\widehat{p \cdot x} = \hat{p} * \hat{x} = 0$.

Here $(x * w)(k) = \sum_{l=0}^{N-1} x(l) w(k-l)$
$\quad$ mod N

We show that $\hat{p}$ solves (1); $\hat{p}(0)=1$, $\hat{p}(k)=0$, $k>s$:

- $$\hat{p}(0)=\sum_{t=0}^{N-1}{}' p(t)=\frac{1}{N}\sum_{t=0}^{N-1}{}' \prod_{k\in\tilde{S}}\left(1-e^{-2\pi i\frac{k}{N}}\cdot e^{2\pi i\frac{t}{N}}\right)$$

$$=\frac{1}{N}\sum_{t=0}^{N-1}{}'\sum_{S'\subset\tilde{S}}{}'(-1)^{\#S'}\cdot\exp\left(2\pi i\frac{t}{N}\right)^{\#S'}\cdot\exp\left(-\frac{2\pi i}{N}\sum_{l\in S'}{}' l\right)$$

$$=\frac{1}{N}\sum_{S'\subset\tilde{S}}{}'(-1)^{\#S'}\exp\left(-\frac{2\pi i}{N}\sum_{l\in S'}{}' l\right)\cdot\underbrace{\sum_{t=0}^{N-1}{}'\exp\left(2\pi i\frac{t}{N}\#S'\right)}_{=\begin{cases}0\ldots \#S'\neq 0\\ N\ldots \#S'=0\end{cases}}$$

$$\overset{S'=\emptyset}{=}\frac{1}{N}(-1)^0\cdot 1\cdot N=1$$

$\ldots\ \hat{p}(k)=0$ for $k>s$ ... in the same way ... supp $\hat{p}\subset\{0,1,\ldots,s\}$.

- From $\hat{p}*\hat{x}=0$ we get

$$\hat{p}*\hat{x}(s)=0:\quad \hat{p}(0)\hat{x}(s)+\hat{p}(1)\hat{x}(s-1)+\ldots+\hat{p}(s)\hat{x}(0)=0$$

$$\hat{p}*\hat{x}(s+1)=0:\quad \hat{p}(0)\hat{x}(s+1)+\hat{p}(1)\hat{x}(s)+\ldots+\hat{p}(s)\hat{x}(1)=0$$

$$\vdots$$

$$\hat{p}*\hat{x}(2s-1)=0:\quad \hat{p}(0)\hat{x}(2s-1)+\hat{p}(1)\hat{x}(2s-2)+\ldots+\hat{p}(s)\hat{x}(s-1)=0$$

... $\hat{p}$ solves (1) ... (1) has (at least) one solution

- If this solution is unique $\Rightarrow\hat{p}=z$

$$S=(\operatorname{supp} z^\vee)^c=(\operatorname{supp} p)^c=(\{0,1,\ldots,N-1\}\setminus\tilde{S})^c=\tilde{S}.$$

Step 2: If (1) has more solutions:

- if z solves (1), $z(0)=1$, $z(k)=0$ for $k>s$, put $q=z^{\vee}$

$$\text{Then } \widehat{q\cdot x}(j)=(\hat{q}*\hat{x})(j)=0 \quad \text{for } j=s,s+1,\dots,2s-1$$
$$=(z*\hat{x})(j)$$

- Then $q\cdot x$ is an s-sparse vector with s vanishing consecutive Fourier coef. $\overset{\text{LEMMA}}{\Longrightarrow}$ $q\cdot x\equiv 0$

$$\Rightarrow q(j)=0 \text{ for } j\in\tilde{S}.$$

$$\Rightarrow \tilde{S}=\operatorname{supp} x\subset\{j: q(j)=0\}=\{j: z^{\vee}(j)=0\}=(\operatorname{supp} z^{\vee})^c=S$$
$$\#\tilde{S}\leq s.$$

The more modern part of Compressed Sensing (~2006)

- If A and y is given, the most natural idea is to look for the most sparse $x \in \mathbb{R}^N$ such that $y = Ax$:

$$x^\# = \operatorname{argmin} \|x\|_0 \text{ s.t. } Ax = y$$

- Recall: We want to design A, such that if $y = Ax$ for some sparse x, then there is (fast?) algorithm, which recovers x from the input (A, y).

- ℓ_0-min. is not convex ... replace by ℓ_1-minimization

$$(P_1) \quad x^\# = \operatorname{argmin} \|z\|_1 \text{ s.t. } Az = y \quad \dots (A,y) \text{ are inputs}$$

Theorem:
- We say that $A \in \mathbb{R}^{m \times N}$ has the Null Space Property of the order s if (NSP_s)

$$\forall v \in \ker A \setminus \{0\} \ \forall S \subset \{1, \dots, N\} : \#S \le s : \|v_S\|_1 < \|v_{S^c}\|_1 .$$

- Every s-sparse $x \in \mathbb{R}^N$ is the unique solution of (P_1) with (A, y), where $y = Ax$ if, and only if, A has (NSP_s).

Proof: "$\Rightarrow$" We assume that every s-sparse x is uniquely found by (P_1).

Let $v \in \ker A \setminus \{0\}$, $S \subset \{1, \dots, N\}$, $\#S \le s$. Then $0 = Av = Av_S + A(v_{S^c})$

... i.e. $A(-v_{S^c}) = Av_S$. But v_S is s-sparse $\Rightarrow$

$$\|v_S\|_1 < \|z\|_1 \text{ for every } z \text{ with } Az = Av_S \ \dots \Rightarrow \|v_S\|_1 < \|v_{S^c}\|_1$$

$\Rightarrow A$ has (NSP_s).

"$\Leftarrow$": We assume that A has (NSP_s), $x\in\mathbb{R}^N$ with $\|x\|_0\leq s$.

We have to show that if $z\in\mathbb{R}^N$ has $Az=Ax$ & $z\neq x$, then $\|x\|_1<\|z\|_1$.

Put $v=x-z$, $S=\operatorname{supp} x$ ... $\underset{v\neq 0}{Av=0}$, $\#S\leq s$

Then $\|x\|_1\leq\|x-z_S\|_1+\|z_S\|_1=\|v_S\|_1+\|z_S\|_1<\|v_{S^c}\|_1+\|z_S\|_1$

$=\|z_{S^c}\|_1+\|z_S\|_1=\|z\|_1$ ∎

Remarks: • NSP_s is defined in a simple language ... kern, $\|.\|_1$, ...
but it is actually difficult to verify if A has (NSP_s).

• If A has NSP_s and $y=Ax$ for an s-sparse $x\in\mathbb{R}^N$, then there is an effective algorithm to find x ... ℓ_1-minimization

Restricted Isometry Property (RIP)

Def: If $A\in\mathbb{R}^{m\times N}$, $1\leq s\leq N$. Then the RIP-constant $\delta_s=\delta_s(A)$ of the order s is the smallest $\delta\geq 0$, such that

$$(1-\delta)\|x\|_2^2\leq\|Ax\|_2^2\leq(1+\delta)\|x\|_2^2 \text{ for every } s\text{-sparse } x\in\mathbb{R}^N.$$

Theorem: Let $A\in\mathbb{R}^{m\times N}$, $s\leq N/2$ and $\delta_{2s}<1/3$. Then A has (NSP_s).

Proof: Let $v\in\operatorname{kern} A\setminus\{0\}$, $S\subset\{1,\dots,N\}$ with $\#S\leq s$.

We show that $(*)$ $\|v_S\|_2\leq\dfrac{\delta_{2s}}{1-\delta_s}\cdot\dfrac{1}{\sqrt{s}}\|v\|_1$

Then $\delta_s\leq\delta_{2s}<1/3$ ... $\|v_S\|_1\leq\sqrt{s}\,\|v_S\|_2\leq\frac{1}{2}\|v\|_1$ ... $2\|v_S\|_1<\|v\|_1$

$\Rightarrow\|v_S\|_1<\|v_{S^c}\|_1$.

Proof of (*) ... it is enough to consider S as the indices of s largest coordinates of v ... we assume that $S=\{1,\dots,s\}$

.. otherwise $S=\{\sigma(1),\dots,\sigma(s)\}$ ... etc.

We set $S_0=S=\{1,\dots,s\}$

$S_1=\{s+1,\dots,2s\}$ ... the s second largest entries of v

$S_2=\{2s+1,\dots,3s\}$

$\vdots$

$|v_1|\geq|v_2|\geq\dots$

By $Av=0$ we have $Av_{S_0}=A(-v_{S_1}-v_{S_2}-\dots)$

$$\Rightarrow \|Av_{S_0}\|_2^2\leq\frac{\|Av_{S_0}\|_2^2}{1-\delta_s}=\frac{1}{1-\delta_s}\langle Av_{S_0},A(-v_{S_1})+A(-v_{S_2})+\dots\rangle$$

$$=\frac{1}{1-\delta_s}\sum_{k\geq 1}\langle Av_{S_0},Av_{S_k}\rangle$$

- If $x,z\in\mathbb{R}^N$, $\|x\|_0,\|z\|_0\leq s$ and the supports of x and z are disjoint, and $\|x\|_2=\|z\|_2=1$

Then $\|x\pm z\|_2^2=2$, $x\pm z$ are $2s$-sparse

$$\Rightarrow (1-\delta_{2s})\cdot 2\leq\|A(x\pm z)\|_2^2\leq 2(1+\delta_{2s})$$

$$\Rightarrow |\langle Ax,Az\rangle|=\frac{1}{4}\left|\|A(x+z)\|_2^2-\|A(x-z)\|_2^2\right|\leq\delta_{2s}.$$

- If x,z are s-sparse, disjoint support & general norms: $\tilde{x}=\frac{x}{\|x\|_2}$, $\tilde{z}=\frac{z}{\|z\|_2}$

$$\Rightarrow|\langle Ax,Az\rangle|\leq\delta_{2s}\cdot\|x\|_2\cdot\|z\|_2.$$

$$\Rightarrow \|v_{S_0}\|_2^2 \leq \frac{1}{1-\delta_s} \sum_{k\geq 1} |\langle A v_{S_0}, A(-v_{S_k})\rangle| \leq \frac{1}{1-\delta_s} \sum_{k\geq 1} \delta_{2s} \cdot \cancel{\|v_{S_0}\|_2} \cdot \|v_{S_k}\|_2$$

$$\Rightarrow \|v_{S_0}\|_2 \leq \frac{\delta_{2s}}{1-\delta_s} \sum_{k\geq 1} \|v_{S_k}\|_2$$

Now $\|v_{S_k}\|_2 = \left(\sum_{j\in S_k} |v_j|^2\right)^{1/2} \leq \left(s \cdot \min_{i\in S_{k-1}} |v_i|^2\right)^{1/2} \leq \sqrt{s} \cdot \min_{i\in S_{k-1}} |v_i| \leq \frac{\|v_{S_{k-1}}\|_1}{\sqrt{s}}$

Finally $\|v_{S_0}\|_2 \leq \frac{\delta_{2s}}{1-\delta_s} \sum_{k\geq 1} \frac{1}{\sqrt{s}} \|v_{S_{k-1}}\|_1 = \frac{\delta_{2s}}{1-\delta_s} \cdot \frac{1}{\sqrt{s}} \|v\|_1 \ldots \Rightarrow (*)$ ∎

Stability: • what if x is not exactly s-sparse but close to it

$\sigma_s(x)_p = \inf\{\|x-y\|_p, y\in\mathbb{R}^N, y \text{ is } s\text{-sparse}\}$ is small...

Robustness: $y = Ax+\varepsilon$, ε... noise

• Robust Null Space Property: $0<\rho<1$, $\tau>0$

$\forall v\in \mathbb{R}^N, \forall \#S\leq s: \|v_S\|_1 \leq \rho\|v_{S^c}\|_1 + \tau\|Av\|_2$

• $(P_{1,\eta})$: $x^\# = \underset{z\in\mathbb{R}^N}{\operatorname{argmin}} \|z\|_1$, s.t. $\|Az-y\|_2 \leq \eta$

Theorem: $A\in\mathbb{R}^{m\times N}$ with $\delta_{2s} < \delta_* = 0.49\ldots$ Then for every $x\in\mathbb{R}^N$ and every $y\in\mathbb{R}^m$ with $\|Ax-y\|_2\leq\eta$, the solution $x^\#$ of $(P_{1,\eta})$ satisfies

$$\|x-x^\#\|_1 \leq C\sigma_s(x)_1 + D\sqrt{s}\,\eta$$

$$\& \quad \|x-x^\#\|_2 \leq \frac{C}{\sqrt{s}}\sigma_s(x)_1 + D\eta. \qquad (C, D \text{ depend on } \delta_{2s})$$

We want to construct matrices $A \in \mathbb{R}^{m \times N}$ with small δ_{2s} and small m

$\Rightarrow$ random constructions.

Lemma • Let $\omega \sim \mathcal{N}(0,1)$ (i.e., with density $p(t) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2})$, $t \in \mathbb{R}$).

Then $\mathbb{E}(e^{\lambda \omega^2}) = \frac{1}{\sqrt{1-2\lambda}}$ for $-\infty < \lambda < 1/2$.

• (2-stability of normal distribution)

Let $m \in \mathbb{N}$, $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$ and let $\omega_1, \dots, \omega_m \sim \mathcal{N}(0,1)$ are indep.

Then $\langle \lambda, \omega \rangle = \lambda_1 \omega_1 + \dots + \lambda_m \omega_m \sim \|\lambda\|_2 \cdot \mathcal{N}(0,1)$

/ Exercises

Concentration of measure ... $\frac{\omega_1^2 + \dots + \omega_m^2}{m}$ concentrates around 1 for m large

Lemma: Let $m \in \mathbb{N}$ and let $\omega_1, \dots, \omega_m \sim \mathcal{N}(0,1)$ be independent. Then

$$\mathbb{P}(\omega_1^2 + \dots + \omega_m^2 \geq (1+\varepsilon)m) \leq \exp\left(-\frac{m}{2}\left[\varepsilon^2/2 - \varepsilon^3/3\right]\right)$$

and

$$\mathbb{P}(\omega_1^2 + \dots + \omega_m^2 \leq (1-\varepsilon)m) \leq \exp\left(-\frac{m}{2}\left[\varepsilon^2/2 - \varepsilon^3/3\right]\right).$$

Proof: 1. inequality, $\beta = 1+\varepsilon$

$$\mathbb{P}(\omega_1^2 + \dots + \omega_m^2 \geq \beta m) \underset{\lambda > 0}{=} \mathbb{P}(\lambda[\omega_1^2 + \dots + \omega_m^2] - \lambda\beta m \geq 0)$$

$$= \mathbb{P}(\exp\{\lambda(\omega_1^2 + \dots + \omega_m^2)\} \cdot \exp(-\lambda\beta m) \geq 1) \leq \mathbb{E}\, e^{-\lambda\beta m} \exp(\lambda(\omega_1^2 + \dots + \omega_m^2))$$

$$= e^{-\lambda\beta m} \cdot (\mathbb{E} e^{\lambda\omega_1^2})^m = e^{-\lambda\beta m} \cdot (1-2\lambda)^{-m/2} \quad \dots \quad 0 < \lambda < 1/2$$

Optimize over $0 < \lambda < 1/2$ ... $\beta = \frac{1}{1-2\lambda}$, ..., $2\lambda = 1 - \frac{1}{\beta}$ $\begin{array}{l} < 1 \\ > 0 \end{array}$

$$\Rightarrow \leq e^{-\beta m(1/2 - \frac{1}{2\beta})} \cdot \beta^{m/2} = e^{-m(\frac{\beta-1}{2})} e^{\frac{m}{2}\ln\beta} = e^{-\frac{\varepsilon m}{2}} \cdot e^{\frac{m}{2}\ln(1+\varepsilon)}$$

- use $\ln(1+t) \le t - \frac{t^2}{2} + \frac{t^3}{3}$, $-1 < t < 1$

$$\Rightarrow \quad \le e^{-\frac{m\varepsilon}{2}} \cdot e^{\frac{m}{2}\left[\varepsilon - \frac{\varepsilon^2}{2} + \frac{\varepsilon^3}{3}\right]} \Rightarrow \blacksquare$$

- "RIP for one point on the sphere"

Theorem: Let $A = \frac{1}{\sqrt{m}} \begin{pmatrix} \omega_{11}, \dots, \omega_{1N} \\ \\ \omega_{m1}, \dots, \omega_{mN} \end{pmatrix}$ with ω_{ij}'s from $\mathcal{N}(0,1)$, indep ("Gaussian matrix")

Let $\|x\|_2 = 1$. Then

$$\mathbb{P}\left(\left|\|Ax\|_2^2 - 1\right| > t\right) \le 2\exp\left(-\frac{m}{2}\left[\frac{t^2}{2} - \frac{t^3}{3}\right]\right) \le 2\exp(-Cmt^2) \text{ for } 0<t<1$$
and some $C>0$.

Proof: $\mathbb{P}\left(\left|\|Ax\|_2^2 - 1\right| > t\right) = \mathbb{P}\left(\left|\sum_{i=1}^{m}\left(\sum_{j=1}^{N} \omega_{ij} x_j\right)^2 - m\right| > tm\right)$

$$= \mathbb{P}\left(\left|\sum_{i=1}^{m} \omega_i^2 - m\right| > tm\right) = \mathbb{P}(\omega_1^2 + \dots + \omega_m^2 \ge (1+t)m)$$

$$+ \mathbb{P}(\omega_1^2 + \dots + \omega_m^2 \le (1-t)m) \le 2\exp\left(-\frac{m}{2}\left[t^2/2 - t^3/3\right]\right) \dots \quad C = 1/12.$$

Remarks: - for general x ... $\tilde{x} = \frac{x}{\|x\|_2}$: $\mathbb{P}\left(\left|\|Ax\|_2^2 - \|x\|_2^2\right| > t\|x\|_2^2\right) \le 2e^{-Cmt^2}$.

- One can prove the theorem also by rotational invariance of Gauss. r.v.

Net argument:

Lemma: Let $t>0$. Then there exists a set $\mathcal{M} \subset \mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : \|x\|_2 = 1\}$, such that i) $\#\mathcal{M} \le (1 + 2/t)^d$

ii) $\forall z \in \mathbb{S}^{d-1}\ \exists x \in \mathcal{M} : \|x - z\|_2 \le t$.

Exercise

Theorem: Let $N \geq m \geq s \geq 1$. Let $0 < \delta, \varepsilon < 1$ with

$$m \geq C\delta^{-2}\left(s\ln\left(\frac{eN}{s}\right) + \ln(2/\varepsilon)\right), \text{ where } C>0 \text{ is a constant.}$$

Then the Gauss matrix $A \in \mathbb{R}^{m\times N}$ has $\mathbb{P}(\delta_s(A) \leq \delta) \geq 1-\varepsilon$.

Proof: <u>Step 1</u>: Let $Z = \{z \in \mathbb{R}^N : \operatorname{supp}(z) \subset \{1,\dots,s\}, \|z\|_2 = 1\}$.

By net-Lemma ($t = 1/4$), there is $\mathcal{M} \subset Z$ with

i) $\#\mathcal{M} \leq 9^s$ ii) $\min\limits_{x\in\mathcal{M}} \|x-z\|_2 \leq 1/4$ for all $z \in Z$.

We show that if $|\|Ax\|_2^2 - 1| \leq \delta/2$ for all $x \in \mathcal{M}$, then

$$|\|Az\|_2^2 - 1| \leq \delta \text{ for all } z \in Z.$$

"bootstrap argument": Let $\gamma > 0$ be the smallest number with

$$|\|Az\|_2^2 - 1| \leq \gamma \text{ for all } z \in Z \dots \text{ we want } \gamma \leq \delta.$$

Then $|\|Au\|_2^2 - \|u\|_2^2| \leq \gamma\|u\|_2^2$ for all $u \in \mathbb{R}^N$, $\operatorname{supp} u \subset \{1,\dots,s\}$.

- Let $u, v \in Z$: Then

$$|\langle Au, Av\rangle - \langle u, v\rangle| = \frac{1}{4}\left|\|A(u+v)\|_2^2 - \|A(u-v)\|_2^2 - (\|u+v\|_2^2 - \|u-v\|_2^2)\right|$$

$$\leq \frac{1}{4}\left|\|A(u+v)\|_2^2 - \|u+v\|_2^2\right| + \frac{1}{4}\left|\|A(u-v)\|_2^2 - \|u-v\|_2^2\right| \leq$$

$$\leq \frac{\gamma}{4}\|u+v\|_2^2 + \frac{\gamma}{4}\|u-v\|_2^2 = \frac{\gamma}{2}(\|u\|_2^2 + \|v\|_2^2) = \gamma$$

- If $u, v \in \mathbb{R}^N$ with $\operatorname{supp} u, \operatorname{supp} v \subset \{1,\dots,s\}$:

$$|\langle Au, Av\rangle - \langle u, v\rangle| \leq \gamma\|u\|_2 \cdot \|v\|_2.$$

By triangle ineq.: $\bullet\ z \in Z \ldots \exists x \in \mathcal{M} : \|z-x\|_2 \le 1/4$

$$\big|\|Az\|_2^2 - 1\big| = \big|\|Ax\|_2^2 - 1 + \langle A(z+x), A(z-x)\rangle - \langle z+x, z-x\rangle\big|$$

$$\le \underbrace{\big|\|Ax\|_2^2 - 1\big|}_{\le \delta/2} + \gamma \underbrace{\|z+x\|_2}_{\le 2} \cdot \underbrace{\|z-x\|_2}_{\le 1/4} \le \frac{\delta}{2} + \frac{\gamma}{2}$$

$\sup\limits_{z \in Z}$ on the left: $\gamma \le \frac{\delta}{2} + \frac{\gamma}{2} \Rightarrow \gamma \le \delta$.

Step 2: The rest is an union bound:

$$\mathbb{P}(\delta_s(A) > \delta) \le \sum_{\substack{S \subset \{1,\dots,N\} \\ \#S = s}} \mathbb{P}\left(\exists z \in \mathbb{R}^N : \operatorname{supp} z \subset S, \|z\|_2 = 1, \big|\|Az\|_2^2 - 1\big| > \delta\right)$$

$$\le \binom{N}{s} \mathbb{P}\left(\exists z \in Z : \big|\|Az\|_2^2 - 1\big| > \delta\right)$$

$$\le \binom{N}{s} \mathbb{P}\left(\exists x \in \mathcal{M} : \big|\|Ax\|_2^2 - 1\big| > \delta/2\right) \le \binom{N}{s} \cdot 9^s \cdot 2e^{-C' m \delta^2}.$$

Exercise $\le \varepsilon$ if m has the condition from the theorem, $C > 0$ large enough. ☒

- Usually (see before) $\delta = 1/3$ is OK.

Optimality of m?

- $m \geq s$ is surely necessary (even when we know the non-zero pos.)
- $m = 2s$ of Prony is not stable
- $m \sim s \ln N$ is OK for Gauss & stable recovery

Question: Is $m \sim s$ impossible for stable recovery?

Lemma: Let $s \leq N$. Then there exist $T_1, \ldots, T_M \subset \{1, \ldots, N\}$ such that

i) $M \geq \left(\frac{N}{4s}\right)^{s/2}$ ii) $\# T_i = s$ for all $i = 1, \ldots, M$

iii) $\#(T_i \cap T_j) < s/2$ for $i \neq j$.

Proof: • Assume $s \leq N/4$ ... otherwise $M = 1$

- Greedy alg.: Take $T_1, T_2, \ldots$ with (ii) and (iii), then we prove (i)
- Let $T \subset \{1, \ldots, N\}$ be arbitrary with $\# T = s$.

The number of sets T' with $\#(T \cap T') \geq s/2$ is

$$\sum_{j=\lceil s/2 \rceil}^{s} \binom{s}{j}\binom{N-s}{s-j} \leq 2^s \cdot \max_{\lceil s/2 \rceil \leq j \leq s} \binom{N-s}{s-j} = 2^s \cdot \binom{N-s}{s - \lceil \frac{s}{2} \rceil} = 2^s \binom{N-s}{\lfloor s/2 \rfloor}.$$

$\Rightarrow$ There is at least $\binom{N}{s} - 2^s \binom{N-s}{\lfloor s/2 \rfloor}$ subsets of $\{1, \ldots, N\}$ with

s elements and $< s/2$ intersection with T.

... If $T_1, \ldots, T_j$ with (ii) and (iii) are already chosen, we can choose

T_{j+1} if $\binom{N}{s} - j \cdot \binom{N-s}{\lfloor s/2 \rfloor} 2^s > 0$. We stop at M, i.e.

$$\binom{N}{s} - 2^s M \binom{N-s}{\lfloor s/2 \rfloor} \leq 0$$

$$\Rightarrow M \geq \frac{\binom{N}{s}}{2^{+s}\binom{N-s}{\lfloor s/2\rfloor}} \geq 2^{-s}\frac{\binom{N}{s}}{\binom{N-\lceil s/2\rceil}{\lfloor s/2\rfloor}} = 2^{-s}\cdot\frac{N!}{s!\cancel{(N-s)!}}\cdot\frac{(\lfloor s/2\rfloor)!\,\cancel{(N-\lceil s/2\rceil-\lfloor s/2\rfloor)!}}{(N-\lceil s/2\rceil)!}$$

$$= 2^{-s}\cdot\frac{N\cdot(N-1)\dots(N-\lceil s/2\rceil+1)}{s\cdot(s-1)\dots\underbrace{(\cancel{s}-\lfloor s/2\rfloor+1)}_{(s-\lceil s/2\rceil+1)}} \geq 2^{-s}\cdot\left(\frac{N}{s}\right)^{\lceil s/2\rceil} \geq 2^{-s}\left(\frac{N}{s}\right)^{s/2}$$

$$= \left(\frac{N}{4s}\right)^{s/2}. \qquad \blacksquare$$

Theorem: Let $s \leq m \leq N$ be integers, let $A \in \mathbb{R}^{m\times N}$ and let

$\Delta : \mathbb{R}^m \to \mathbb{R}^N$ be an arbitrary mapping with

$$\|x - \Delta(Ax)\|_2 \leq \frac{C\sigma_s(x)_1}{\sqrt{s}} \text{ for all } x \in \mathbb{R}^N.$$

Then $m \geq C' s \ln\left(\frac{eN}{s}\right)$.

Proof: We may assume $C \geq 1$, $s \leq N/8$.

By Lemma: $T_1, \dots, T_M$ ... put $x_i := \chi_{T_i}\cdot\frac{1}{\sqrt{s}}$

Then: $\|x_i\|_2 = 1$, $\|x_i - x_j\|_2 > 1$, $\|x_i\|_1 = \sqrt{s}$.

Put $B := \left\{z \in \mathbb{R}^N : \|z\|_1 \leq \frac{\sqrt{s}}{4C} \;\&\; \|z\|_2 \leq 1/4\right\}$.

Then $x_i \in 4CB$. We claim that $A(x_i + B)$ are mutually disjoint.

If not: $\exists i \neq j\ \exists z, z' \in B : A(x_i+z) = A(x_j+z') \dots \;\&\; \Delta(A(x_i+z)) = \Delta(A(x_j+z'))$

$$\Rightarrow 1 < \|x_i - x_j\|_2 = \|[x_i + z - \Delta(A(x_i+z))] - [x_j + z' - \Delta(A(x_j+z'))] - z + z'\|_2$$

$$\leq \|x_i + z - \Delta(A(x_i+z))\|_2 + \|x_j + z' - \Delta(A(x_j+z'))\|_2 + \|z\|_2 + \|z'\|_2$$

$$\leq \frac{C\sigma_s(x_i+z)_1}{\sqrt{s}} + \frac{C\sigma_s(x_j+z')_1}{\sqrt{s}} + 1/4 + 1/4 \overset{x_i, x_j\ s\text{-sparse}}{\leq} \frac{C\|z\|_1}{\sqrt{s}} + \frac{C\|z'\|_1}{\sqrt{s}} + 1/2 \leq 1.$$

Finally, $A(x_i+B) \subset A((4C+1)B)$ and we compare volume

• $d \leq m$ is the rank of A .. $V = vol_d(A(B))$

$$\Rightarrow \underbrace{\sum_{j=1}^{M} vol(A(x_j+B))}_{\substack{\text{\rotatebox{90}{$\leq$}} \\ M\, vol(A(B)) \\ \text{\rotatebox{90}{$\leq$}} \\ \left(\frac{N}{4s}\right)^{s/2} \cdot V}} \leq \underbrace{vol((4C+1)A(B))}_{= (4C+1)^d \cdot V \leq (4C+1)^m \cdot V}$$

$$\Rightarrow \left(\frac{N}{4s}\right)^{s/2} \leq (4C+1)^m$$

$$\Rightarrow \frac{s}{2} \ln\left(\frac{N}{4s}\right) \leq m \ln(4C+1) \underset{s \leq \frac{N}{8}}{\Rightarrow} \blacksquare$$