Compressed Sensing (Day 3)

Sadashige Ishida & Raimundo Saona

June 9, 2020

Review of the previous lectures

Previously, we learned the concept of compressed sensing:

In a real situation, we want to know the information of a signal from measurements. Intuitively, we need all the measurements to obtain the signal. However, we can likely recover the signal from partial measurements via ℓ_0 or ℓ_1 minimization problems.

A fundamental question is:

When is a recovery possible?

Main theorem review

Our goal is to prove this probabilistic result.

Theorem 3.1 (Candes, Romberg, and Tao [1])

Let $f \in \mathbb{C}^N$ be a signal, and Ω be the random set, and M be an accuracy parameter. If f is supported on $T \subset \{0, \dots, N-1\}$ and,

$$\mathbb{E}(|\Omega|) \geq |T| \log N / \alpha(M), \tag{3.1}$$

then, f can be exactly recovered from Ω and $\hat{f}|_\Omega$ with probability at least $1 - O(N^{-M})$ as the unique minimizer of the ℓ_1-problem,

$$\min_{g \in \mathbb{C}^N} \|g\|_1 := \sum_t |g(t)|, \quad \hat{g}|_\Omega = \hat{f}|_\Omega. \tag{3.2}$$

The dual problem

Last time, we learned a dual problem:

Theorem 3.2

Let $\Omega \subset \{0, \ldots, N-1\}$. For a signal $f \in \mathbb{C}^N$ with $\operatorname{supp}(f) = T$*. Suppose $\mathcal{F}_{T\to\Omega}$ is injective and there exists a polynomial P s.t.*

$$P(t) = \operatorname{sign}(f)(t) := f(t)/|f(t)| \quad \text{for } t \in T, \tag{3.3}$$

$$|P(t)| < 1 \quad \text{for } t \notin T, \tag{3.4}$$

and

$$\hat{P}(\omega) = 0 \quad \text{for } \omega \notin \Omega. \tag{3.5}$$

Then, there exists a unique minimizer of the ℓ_1-problem, which equals to f.

Load map to the proof

We constructed a polynomial P as a candidate of the one stated in the theorem,

$$P := (\imath - \frac{1}{|\Omega|}H)(\imath^*\imath - \frac{1}{|\Omega|}\imath^*H)^{-1}\imath^*\mathrm{sign}(f) \tag{3.6}$$

where $H : \mathbb{C}^T \to \mathbb{C}^N$ is defined by

$$Hf(t) := -\sum_{\omega\in\Omega} \sum_{t'\in T : t'\neq t} e^{-i\omega(t-t')} \tag{3.7}$$

and $\imath : \mathbb{C}^T \to \mathbb{C}^N,\ \imath^* : \mathbb{C}^N \to \mathbb{C}^T$ are the inclusion and the restriction.

Load map to the proof

We checked

$$P = \mathrm{sign}(f) \quad \text{on } T, \tag{3.8}$$

$$\hat{P} = 0 \quad \text{on } \Omega^c, \tag{3.9}$$

and that the invertiability of

$$\imath^*\imath - \frac{1}{|\Omega|}\imath^* H = \frac{1}{|\Omega|}\mathcal{F}^*_{T\to\Omega}\,\mathcal{F}_{T\to\Omega} \tag{3.10}$$

implies the injectivity of $\mathcal{F}_{T\to\Omega}$.

Load map to the proof

Our task is now to verify

1. Invertiability of $\imath^*\imath - \frac{1}{|\Omega|}\imath^* H$
2. $|P| < 1$ on T^c

with probability at least $1 - O(N^{-M})$.

Today's contents

We prove 1 and some results for 2.

Matrix norms

Let M be a $n \times m$ complex valued matrix. We use the following norms.

Infinity norm
$$\|M\|_\infty := \max_i \sum_j |M_{i,j}|, \tag{3.11}$$

Operator norm
$$\|M\| := \sup_{\|x\|_2=1} \|Mx\|_2, \tag{3.12}$$

Frobenius norm
$$\|M\|_F := \sqrt{\sum_{i,j} |M_{i,j}|^2}. \tag{3.13}$$

Note all matrix norms are equivalent. Namely, convergence in one norm means convergence in all the norms. The space of matrices equipped with these norms is complete (Banach).

Properties of $\|\cdot\|_\infty$

- $\|A+B\|_\infty \le \|A\|_\infty + \|B\|_\infty$
- $\|AB\|_\infty \le \|A\|_\infty \|B\|_\infty$
- $\|A\|_\infty = \sup_{\|x\|_\infty=1} \|Ax\|_\infty$

Properties of $\|\cdot\|$

- $\|A + B\| \leq \|A\| + \|B\|$
- $\|AB\| \leq \|A\| \|B\|$
- $\|A^*\| = \|A\|$
- $\|A^2\| = \|A\|^2$ if A is self adjoint
- $\max_i |\lambda_i| \leq \|A\|$

Properties of $\|\cdot\|_F$

- $\|A\|_F = \sqrt{\mathrm{Tr}(A^*A)}$
- $\|A+B\|_F \leq \|A\|_F + \|B\|_F$ (HW)
- $\|AB\|_F \leq \|A\|_F \|B\|_F$ (HW)
- $\|A\| \leq \|A\|_F$
- $\|A\|_\infty \leq \#\mathrm{col}(A) \|A\|_F$

Properties of matrix norms

Let A be a self-adjoint matrix. Then

$$\left\|A^2\right\| = \|A\|^2 . \tag{3.14}$$

Proof It is immediate that,

$$\left\|A^2\right\| \leq \|A\| \, \|A\| = \|A\|^2 . \tag{3.15}$$

The other inequality is by the Cauchy-Schwartz inequality,

$$\|A\|^2 = \sup_{\|x\|_2=1} \|Ax\|_2^2 = \sup_{\|x\|_2=1} |\langle Ax, Ax\rangle| = \sup_{\|x\|_2=1} |\langle A^*Ax, x\rangle| \tag{3.16}$$

$$\leq \sup_{\|x\|_2=1} \|A^*Ax\|_2\|x\|_2 = \|A^2\|. \tag{3.17}$$

Properties of matrix norms

For a matrix A,

$$\|A\| \leq \|A\|_F \,. \tag{3.18}$$

Proof For any vector x, $\|x\|_2 = \|x\|_F$ by definition. Therefore we have,

$$\|A\| = \sup_{\|x\|_2=1} \|Ax\|_2 = \sup_{\|x\|_F=1} \|Ax\|_F \leq \sup_{\|x\|_F=1} \|A\|_F \|x\|_F = \|A\|_F \,. \tag{3.19}$$

Our strategy

To prove the invertiability of $\imath^*\imath - \frac{1}{|\Omega|}\imath^* H$, we show

$$\left\| \frac{1}{|\Omega|}\imath^* H \right\| < 1. \tag{3.20}$$

with a higher probability than we need. To see why, we know that

$$\max_k |\lambda_k| \leq \|M\| \tag{3.21}$$

for any matrix M.[1] Then the eigenvalues of $I - M$ is bounded by $1 - \max_k |\lambda_k|$ and $1 + \max_k |\lambda_k|$.

[1] In fact the equality holds in our case since $\imath^* H$ is self adjoint.

Neumann Series

Moreover, we can explicitly find $(\imath^*\imath - \frac{1}{|\Omega|}\imath^* H)^{-1}$ as a Neumann series, which will be useful when we show the boundeddness of $|P|$.

Proposition 3.1 (The Neumann Series)

Let M be a square matrix. Then $\|M\| < 1$ implies that $\sum_{n=0}^{\infty} M^n$ converges and equals to $(I - M)^{-1}$.

Proof We first show the existence of the limit. Due to the completeness of the space, it suffices to prove that the partial sum

$$S_n := \sum_{i=0}^{n} M^i \qquad (3.22)$$

is a Cauchy series in $\|\cdot\|$. For k and l with $k > l$,

$$\begin{aligned}\|S_k - S_l\| &= \left\| \sum_{i=l+1}^{k} M^i \right\| \leq \sum_{i=l+1}^{k} \left\|M^i\right\| \leq \sum_{i=l+1}^{k} \|M\|^i \\ &= \|M\|^{l+1} \frac{(1 - \|M\|^{k-l})}{1 - \|M\|} \to 0 \quad \text{as } l, k \to \infty.\end{aligned}$$

Due to the equivalence of norms, this implies the convergence in $\max \| \cdot \|$ norm i.e. all entries converge. Hence $S := \lim_{n\to\infty} S_n$ exists.

Neumann Series

We then show $S = (I - M)^{-1}$. First,

$$MS_n = \sum_{i=1}^{n+1} M^i = \sum_{i=0}^{n+1} M^i - I = S_{n+1} - I. \tag{3.23}$$

From this, it follows

$$\begin{aligned} \|MS - (S - I)\| &= \|MS - MS_n + (S_{n+1} - I) - (S - I)\| && (3.24) \\ &\leq \|MS - MS_n\| + \|(S_{n+1} - I) - (S - I)\| && (3.25) \\ &\leq \|M\| \, \|S - S_n\| + \|S_n - S\| \to 0. && (3.26) \end{aligned}$$

Hence, we obtained

$$I = S - MS = S(I - M). \tag{3.27}$$

By doing the same with $S_n M$, we get $I = (I - M)S$.

Strategy review

We hereby denote $H_0 := \imath^* H$ for simplicity.

Remark

In order to prove the invertiability of $\imath^*\imath - \frac{1}{|\Omega|} H_0$ for a probability at least $1 - O(N^{-M})$, our goal is to show $\|H_0\| < |\Omega|$ for an equally high probability.

Decomposition of events

We would like to attain

$$\mathbb{P}\left(\|H_0\| \geq |\Omega|\right) \leq O(N^{-M}). \tag{3.28}$$

For this purpose, we consider a decomposition of events

$$\begin{aligned}
\{\|H_0\| \geq |\Omega|\} =& \{(1-\epsilon_M)\tau N \geq \|H_0\| \geq |\Omega|\} && (3.29)\\
& \cup \{\|H_0\| \geq (1-\epsilon_M)\tau N \geq |\Omega|\} && (3.30)\\
& \cup \{\|H_0\| \geq |\Omega| \geq (1-\epsilon_M)\tau N\} && (3.31)
\end{aligned}$$

where

$$\epsilon_M := \frac{2M \log N}{\tau N} \tag{3.32}$$

Note $\tau N = \mathbb{E}\left(|\Omega|\right)$ since Ω is defined using Binominal distribution. Our goal is to see the probability of each event set is bounded to $O(N^{-M})$.

Decomposition of events

Regarding the first set, we note that

$$\{(1-\epsilon_M)\tau N \geq \|H_0\| \geq |\Omega|\} \subset \{(1-\epsilon_M)\tau N \geq |\Omega|\} \tag{3.33}$$

and we already obtained

$$\mathbb{P}\left((1-\epsilon_M)\tau N \geq |\Omega|\right) \leq O(N^{-M}) \tag{3.34}$$

in the last lecture.

Decomposition of events

Regarding the second and the thrid set,

$$\{\|H_0\| \geq (1-\epsilon_M)\tau N \geq |\Omega|\}, \quad \{\|H_0\| \geq |\Omega| \geq (1-\epsilon_M)\tau N\}, \tag{3.35}$$

their union is contained in

$$\{\|H_0\| \geq (1-\epsilon_M)\tau N\}. \tag{3.36}$$

We aim to see the probability measure of this set is bounded by $O(N^{-M})$.

Load map to upper bound

To obtain an upper bound of $\mathbb{P}(\|H_0\| \geq (1-\epsilon_M)\tau N)$, we concern $\mathbb{E}\left(\mathrm{Tr}(H_0^{2n})\right)$. This is because

$$\mathbb{P}(\|H_0\| \geq (1-\epsilon_M)\tau N) = \mathbb{P}\left(\|H_0\|^{2n} \geq ((1-\epsilon_M)(\tau N))^{2n}\right) \tag{3.37}$$

and it follows from the self-adjointness of H_0 that

$$\|H_0\|^{2n} = \|H_0^n\|^2 \leq \|H_0^n\|_F^2 = \mathrm{Tr}(H_0^{n*}H_0^n) = \mathrm{Tr}(H_0^{2n}). \tag{3.38}$$

Then the Markov inequality asserts for any $C > 0$,

$$\mathbb{P}\left(\mathrm{Tr}(H_0^{2n}) \geq C\right) \leq \frac{\mathbb{E}\left(\mathrm{Tr}(H_0^{2n})\right)}{C}. \tag{3.39}$$

As such, we aim to establish a good bound of $\mathbb{E}\left(\mathrm{Tr}(H_0^{2n})\right)$.

Probabilistic estimate of $\|H_0\|$

We use a result from random matrix theory.

Lemma 3.3

Set $c_\tau := e\log((1-\tau)/\tau)$ *and* $\phi := (1+\sqrt{5})/2$. *Then*

$$\mathbb{E}\left(\mathrm{Tr}(H_0^{2n})\right) \le n\phi^{2n}\max(a_n, b_n) \tag{3.40}$$

where

$$a_n = (2n-1)^{2n} c_\tau^{-2n-1} N|T|^{2n}, \quad b_n = \frac{(2n)!}{n!2^n}\left(\frac{\tau}{1-\tau}\right)^n N^n|T|^{n+1}. \tag{3.41}$$

In a certain setting, this bound can be simplified so that it does not contain max. We now work on that.

Probabilistic estimate of $\|H_0\|$

Now we assume the following. From now on, the results require these assumptions.

Working assumptions

Let constants ϕ, γ be

$$\phi := \frac{1+\sqrt{5}}{2}, \quad \gamma := \sqrt{\frac{2\phi^2}{1-\tau}}, \tag{3.42}$$

and variables $0 < \alpha < 1$, $\alpha_M := \alpha(1-\epsilon_M)$, and $n \in \mathbb{N}$. For the time being, we assume T satisfies,

$$|T| \leq \frac{\alpha_M^2}{\gamma^2}\frac{\tau N}{n}. \tag{3.43}$$

Recall the upper bound of $|T|$ was required in the statement of the main theorem. We will later specify α and n, which gives the necessary bound for $|T|$.

Probabilistic estimate of $\|H_0\|$

We further assume

Working assumptions

$$0 < \tau \leq 0.44. \tag{3.44}$$

Therefore, our proof requires this bound. This may look a bit odd since the main theorem allows arbitrary τ in $(0, 1]$. This, however, works since if

$$\mathbb{P}\left(\|H_0\| \geq (1 - \epsilon_M)\tau N\right) < O(N^{-M}) \tag{3.45}$$

for $\tau = 0.44$, it holds for $\tau \geq 0.44$ as well.

We simplify the bound of a_n. The next lemma can be proved via a numerical consideration.

Lemma 3.4

With the above setting, we have the bound,

$$a_n := (2n-1)^{2n} c_\tau^{-2n-1} N|T|^{2n} \leq 2^{2n+1} e^{-n} n^n \left(\frac{\tau}{1-\tau}\right)^n N^n |T|^n. \quad (3.46)$$

The next step is to establish a bound for $\max(a_n, b_n)$.

Probabilistic estimate of $\|H_0\|$

For this purpose, we use an approximation.

Lemma 3.5 (The Stirling approximation)

As integer n goes to ∞, it holds asymptotically

$$n! \sim \sqrt{2\pi}\sqrt{n}\left(\frac{n}{e}\right)^n. \tag{3.47}$$

Moreover, for any $n \in \mathbb{N}$, the bound

$$\sqrt{2\pi}\sqrt{n}\left(\frac{n}{e}\right)^n \leq n! \leq e\sqrt{n}\left(\frac{n}{e}\right)^n \tag{3.48}$$

holds.

We note that

$$2.50 < \sqrt{2\pi} < \frac{n!}{\sqrt{n}\left(\frac{n}{e}\right)^n} < e < 2.80. \tag{3.49}$$

Probabilistic estimate of $\|H_0\|$

Recall b_n has a coefficient $\frac{(2n)!}{n!2^n}$. Using the Striling approximation, we obtain

$$\frac{(2n)!}{n!2^n} \leq \frac{1}{2^n}\frac{2.80\sqrt{2n}\left(\frac{2n}{e}\right)^{2n}}{2.50\sqrt{n}\left(\frac{n}{e}\right)^{n}} = \sqrt{2}\frac{2.80}{2.50}2^n e^{-n} n^n \leq 2^{n+1}e^{-n}n^n. \quad (3.50)$$

Then the both of

$$a_n \leq 2^{2n+1}e^{-n}n^n\left(\frac{\tau}{1-\tau}\right)^n N^n|T|^n \text{ and } b_n = \frac{(2n)!}{n!2^n}\left(\frac{\tau}{1-\tau}\right)^n N^n|T|^{n+1}$$

are bounded by

$$2^{2n+1}e^{-n}n^n\left(\frac{\tau}{1-\tau}\right)^n N^n|T|^{n+1}. \quad (3.51)$$

Probabilistic estimate of $\|H_0\|$

We have now,

$$\mathbb{E}\left(\mathrm{Tr}(H_0^{2n})\right) \leq n\phi^{2n}\max(a_n, b_n) \tag{3.52}$$

$$\leq n\phi^{2n}2^{2n+1}e^{-n}n^n\left(\frac{\tau}{1-\tau}\right)^n N^n|T|^{n+1} \tag{3.53}$$

$$= 2e^{-n}\gamma^{2n}n^{n+1}(\tau N)^n|T|^{n+1}. \tag{3.54}$$

Probabilistic estimate of $\|H_0\|$

Summarizing the arguments, we have now the following result.

Theorem 3.6

Let $0 < \alpha < 1$, $\alpha_M := \alpha(1 - \epsilon_M)$, $n \in \mathbb{N}$, $\tau \leq 0.44$ *and* ϕ, γ *as defined before. If* T *satisfies,*

$$|T| \leq \frac{\alpha_M^2}{\gamma^2} \frac{\tau N}{n}, \tag{3.55}$$

then it holds,

$$\mathbb{E}\left(\mathrm{Tr}(H_0^{2n})\right) \leq 2e^{-n}\gamma^{2n} n^{n+1} (\tau N)^n |T|^{n+1}. \tag{3.56}$$

Asymptotic estimate of $\|H_0\|$

The following corollary gives an asymptotic estimate of $\|H_0\|$.

Corollary 3.7

Suppose

$$|T| \leq \frac{\tau N}{\log(\tau N)}. \tag{3.57}$$

Then, for any $\epsilon > 0$, *it holds*

$$\mathbb{E}\left(\|H_0\|\right) \leq \gamma\sqrt{|T|\tau N \log |T|}(1 + o(1)), \quad \text{as } |T| \to \infty \tag{3.58}$$

and

$$\mathbb{P}\left(\|H_0\| > (1+\epsilon)\gamma\sqrt{|T|\tau N \log |T|}\right) \to 0, \quad \text{as } |T| \to \infty. \tag{3.59}$$

Asymptotic estimate of $\|H_0\|$

Proof To gain a bound of $\mathbb{E}(\|H_0\|)$, we would like to leverage our estimate of $\mathbb{E}\left(\mathrm{Tr}(H^{2n})\right)$. Using the self-adjointness of H_0, we get for any $n \in \mathbb{N}$,

$$\|H_0\|^{2n} = \|H_0^n\|^2 \leq \|H_0^n\|_F^2 = \mathrm{Tr}(H_0^{n*} H_0^n) = \mathrm{Tr}(H_0^{2n}). \tag{3.60}$$

We then set $n := \lceil \log |T| \rceil$ to get

$$e^{-n} n^n |T| \leq \lceil \log |T| \rceil^n \tag{3.61}$$

where $\lceil \cdot \rceil$ is the ceil function.
We thus obtain

$$\begin{aligned}
\mathbb{E}(\|H_0\|) &= ((\mathbb{E}(\|H_0\|))^{2n})^{1/2n} \leq (\mathbb{E}(\|H_0\|^{2n}))^{1/2n} \leq (\mathbb{E}(\mathrm{Tr}(H_0^{2n})))^{1/2n} \\
&\leq (2n\gamma^{2n} e^{-n} n^n |T|^{n+1} (\tau N)^n)^{1/2n} && (3.62) \\
&\leq (2n)^{1/2n} \gamma \sqrt{|T| \tau N} \sqrt{\lceil \log |T| \rceil} && (3.63) \\
&= \gamma \sqrt{|T| \tau N} \sqrt{\log |T|} (1 + o(1)) && (3.64)
\end{aligned}$$

Asymptotic estimate of $\|H_0\|$

From these computations and the Markov inequality, we get for any ϵ,

$$
\begin{aligned}
&\mathbb{P}\left(\|H_0\| > (1+\epsilon)\gamma\sqrt{|T|\tau N \log|T|}\right) \\
&= \mathbb{P}\left(\|H_0\|^{2n} > (1+\epsilon)^{2n}\gamma^{2n}|T|^n(\tau N \log|T|)^n\right) \\
&\leq \mathbb{E}\left(\|H_0\|^{2n}\right) / (1+\epsilon)^{2n}\gamma^{2n}|T|^n(\tau N \log|T|)^n \\
&\leq \frac{(2n)^{1/2n}\gamma^{2n}|T|^n(\tau N \lceil \log|T| \rceil)^n}{(1+\epsilon)^{2n}\gamma^{2n}|T|^n(\tau N \log|T|)^n} \\
&= \frac{(2n)^{1/2n}}{(1+\epsilon)^n}\left(\frac{\lceil \log|T| \rceil}{(1+\epsilon)\log|T|}\right)^n \to 0 \quad \text{as } n \to \infty,
\end{aligned}
$$

which completes the proof.

Probabilistic estimate of $\|H_0\|$

Let's return to the main goal. By the Markov inequality, we have

$$\mathbb{P}\left(\|H_0^n\|_F \geq \alpha_M^n(\tau N)^n\right) \leq \frac{\mathbb{E}\left(\|H_0^n\|_F^2\right)}{\alpha_M^{2n}(\tau N)^{2n}} = \frac{\mathbb{E}\left(\mathrm{Tr}(H_0^{2n})\right)}{\alpha_M^{2n}(\tau N)^{2n}}. \tag{3.65}$$

Applying the inequality (Theorem 3.6)

$$\mathbb{E}\left(\mathrm{Tr}(H_0^{2n})\right) \leq 2e^{-n}\gamma^{2n}n^{n+1}(\tau N)^n|T|^{n+1}, \tag{3.66}$$

and the assumption $\frac{\alpha_M^2\tau N}{\gamma^2 n} \leq |T|$, we have now

$$\begin{aligned}\mathbb{P}\left(\|H_0^n\|_F \geq \alpha_M^n(\tau N)^n\right) &\leq (2n)e^{-n}\left(\frac{n\gamma^2}{\alpha^2\tau N}\right)^n |T|^{n+1} \qquad (3.67)\\ &\leq 2e^{-n}\tau N\frac{\alpha_M^2}{\gamma^2}. \qquad (3.68)\end{aligned}$$

Probabilistic estimate of $\|H_0\|$

Using the self adjointness of H_0 and $\|\cdot\| \leq \|\cdot\|_F$, we get

$$\begin{aligned}
\mathbb{P}\left(\|H_0\| \geq \alpha_M \tau N\right) &= \mathbb{P}\left(\|H_0\|^n \geq \alpha_M^n (\tau N)^n\right) && (3.69)\\
&= \mathbb{P}\left(\|H_0^n\| \geq \alpha_M^n (\tau N)^n\right) && (3.70)\\
&\leq \mathbb{P}\left(\|H_0^n\|_F \geq \alpha_M^n (\tau N)^n\right) && (3.71)\\
&\leq 2e^{-n}\tau N \frac{\alpha_M^2}{\gamma^2} \ \leq \ 2e^{-n}\tau N \frac{1}{\gamma^2}. && (3.72)
\end{aligned}$$

Finally, by defining $n := \lceil (M+1) \log N \rceil$, we have

$$\frac{2}{\gamma^2} e^{-n} \tau N \leq \frac{2}{\gamma^2} \frac{\tau N}{N} N^{-M} \leq \frac{2}{\gamma^2} N^{-M}. \tag{3.73}$$

Hence, it holds

$$\mathbb{P}\left(\|H_0\| \geq (1-\epsilon_M)\tau N\right) \leq \mathbb{P}\left(\|H_0\| \geq \alpha_M \tau N\right) \leq \frac{2}{\gamma^2} N^{-M}. \tag{3.74}$$

Probabilistic estimate of $\|H_0\|$

Thus, we have

$$\begin{aligned}\mathbb{P}\left(\|H_0\| \geq |\Omega|\right) &\leq \mathbb{P}\left(|\Omega| \leq (1-\epsilon_M)\tau N\right) + \mathbb{P}\left(\|H_0\| \geq (1-\epsilon_M)\tau N\right) \\ &\leq (\frac{2}{\gamma^2}+1)N^{-M}. \qquad (3.75)\end{aligned}$$

Hence, we have proved $\mathbb{P}\left(\|H_0\| < |\Omega|\right)$ at least probability $1 - O(N^{-M})$, namely $(\imath^*\imath - \frac{1}{|\Omega|}\imath^* H)$ is invertiable.

Invertiability of $\imath^*\imath - \frac{1}{|\Omega|}\imath^* H$

We have now established a conclusion.

Theorem 3.8

Let $0 < \alpha < 1$, $\alpha_M := \alpha(1 - \epsilon_M)$, *If* T *obeys,*

$$|T| \leq \frac{\alpha_M^2}{\gamma^2} \frac{\tau N}{\lceil (M+1) \log N \rceil}, \tag{3.76}$$

then $(\imath^*\imath - \frac{1}{|\Omega|}\imath^* H)$ *is invertiable at least probability* $1 - O(N^{-M})$.

At this point, this result holds for any α as far as T satisfies the above inequality. We will later specify α when guaranteeing $|P| < 1$ on T^c.

Magnitude of P on T^c

In the next lecture, we will prove

$$|P(t)| = \left|(\imath - \frac{1}{|\Omega|}H)(\imath^*\imath - \frac{1}{|\Omega|}\imath^*H)^{-1}\imath^*\mathrm{sign}(f)(t)\right| < 1 \qquad (3.77)$$

on T^c.
Today, we will give an estimate as a warm-up.

Truncated Neumann series

For a square matrix M, we learned the Neumann series expression,

$$(I - M)^{-1} = M^0 + M^1 + M^2 + \cdots . \tag{3.78}$$

The below also holds for any n,

$$(I - M^n)^{-1} I = M^0 + M^n + M^{2n} + \cdots \tag{3.79}$$

$$(I - M^n)^{-1} M^1 = M^1 + M^{n+1} + M^{2n+1} + \cdots \tag{3.80}$$

$$\vdots \tag{3.81}$$

$$(I - M^n)^{-1} M^{n-1} = M^{n-1} + M^{2n-1} + M^{3n-1} + \cdots \tag{3.82}$$

By summing up 3.79 to 3.82, we have

$$(I - M)^{-1} = (I - M^n)^{-1}(1 + M + M^2 + \cdots + M^{n-1}). \tag{3.83}$$

The reminder term R_n

Using this expression, we have

$$
\begin{aligned}
&(\imath^*\imath - \frac{1}{|\Omega|}\imath^* H)^{-1} && (3.84)\\
&= \left(\imath^*\imath - \frac{1}{|\Omega|^n}(\imath^* H)^n\right)^{-1} \sum_{m=0}^{n-1} \frac{1}{|\Omega|^m}(\imath^* H)^m && (3.85)\\
&= \left(\imath^*\imath + \sum_{m=1}^{\infty}\left(\frac{1}{|\Omega|^n}(\imath^* H)^n\right)^m\right) \sum_{m=0}^{n-1} \frac{1}{|\Omega|^m}(\imath^* H)^m && (3.86)\\
&=: (\imath^*\imath + R_n) \sum_{m=0}^{n-1} \frac{1}{|\Omega|^m}(\imath^* H)^m && (3.87)
\end{aligned}
$$

Today, we will estimate that R_n is small in $\|\cdot\|_\infty$ norm.

The reminder term R_n

We first evaluate R_n by the Frobenius norm. Suppose that

$$\|\imath^* H\|_F \leq \alpha|\Omega| \tag{3.88}$$

with a high probability (which we will see in the next lecture). Then,

$$\begin{aligned}
\|R_n\|_F = \left\| \sum_{m=1}^{\infty} \frac{1}{|\Omega|^{mn}} (\imath^* H)^{mn} \right\|_F &\leq \sum_{m=1}^{\infty} \frac{1}{|\Omega|^{mn}} \|(\imath^* H)^{mn}\|_F \\
&\leq \sum_{m=1}^{\infty} \frac{1}{|\Omega|^{mn}} \|\imath^* H\|_F^{mn} \\
&\leq \sum_{m=1}^{\infty} \alpha^{mn} \\
&= \frac{\alpha^n}{1-\alpha^n}.
\end{aligned}$$

Inequality between $\|\cdot\|_\infty$ and $\|\cdot\|_F$

Finally, we use a relation

$$\|M\|_\infty \leq \|M\|_F \sqrt{\#\mathrm{col}(M)} \tag{3.89}$$

for any matrix M.

Proof:

$$\|M\|_\infty := \max_i \sum_j |M_{i,j}| \tag{3.90}$$

$$= \max_i \left(|M_{i,1}|, \cdots, |M_{i,\#\mathrm{col}(M)}|\right) \cdot (1, \ldots, 1) \tag{3.91}$$

$$\leq \max_i \left\|\left(|M_{i,1}|, \cdots, |M_{i,\#\mathrm{col}(M)}|\right)\right\| \cdot \|(1, \ldots, 1)\| \tag{3.92}$$

$$= \max_i \sqrt{\sum_j |M_{i,j}|^2} \sqrt{\#\mathrm{col}(M)} \tag{3.93}$$

$$\leq \sqrt{\sum_{i,j} |M_{i,j}|^2} \sqrt{\#\mathrm{col}(M)} \tag{3.94}$$

$$=: \|M\|_F \sqrt{\#\mathrm{col}(M)}. \tag{3.95}$$

$\|R_n\|_\infty$ is small

We have now,

$$\|R_n\|_\infty \leq \|R_n\|_F \sqrt{\#\text{col}(R_n)} \leq \frac{\alpha^n}{1-\alpha^n}\sqrt{|T|}, \tag{3.96}$$

which approaches to 0 as n increases since $0 < \alpha < 1$.

Summary

Today, we showed:

- With probability at least $1 - O(N^{-M})$, $\imath^*\imath - \frac{1}{\Omega}\imath^* H$ is invertiable, namely $\mathcal{F}_{T\to\Omega}$ is injective.
- $R_n := \sum_{m=1}^{\infty} \frac{1}{|\Omega|^{mn}}(\imath^* H)^{mn}$ is small in $\|\cdot\|_\infty$ under some assumption.

In the next lecture, we will:

- Prove $|P| < 1$ on T^c, which completes the proof of the main theorem.
- See related topics.

E. Candes, J. Romberg, and T. Tao.
Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, 2004.