Sparsity in Time-Frequency Representations by Gotz E. Pfander and Holger Rauhut

R Ramu Naidu

IISc, SPC lab

April 21, 2017

Organization

- Basics of Gabor system
- Basics of Compressed Sensing
- Problem statement
- Main results

Basics of Gabor system

Definition

A translation operator T_k is a mapping from $\mathbb{C}^n \to \mathbb{C}^n$ such that

$$(T_k h)_q = h_{k+q(mod\ n)}; \forall h \in \mathbb{C}^n.$$

Definition

A modulation operator M_l is a mapping from $\mathbb{C}^n \to \mathbb{C}^n$ such that

$$(M_l h)_q = e^{\frac{2\pi ilq}{n}} h_q; \forall h \in \mathbb{C}^n.$$

Definition

The composition of these two operators is called as a time-frequency shift operator. It is a mapping from $\mathbb{C}^n \to \mathbb{C}^n$ such that

$$\Pi(\lambda) = M_l o T_k, \text{ where } \lambda = (k, l).$$

Basics of Gabor system

- The system $\{\Pi(\lambda); \lambda \in \mathbb{Z}_n \times \mathbb{Z}_n\}$ of all time frequency shifts forms a basis for $\mathbb{C}^{n\times n}$.

Definition

For a non zero vector g, the so called window, the set $\{\Pi(\lambda)g; \lambda \in \mathbb{Z}_n \times \mathbb{Z}_n\}$ is called a Gabor system.

- The Gabor system is a tight frame in $\mathbb{C}^n$, whenever $g \neq 0$.

Definition

The matrix $\Psi_g \in \mathbb{C}^{n\times n^2}$ whose columns are the members of a Gabor system is referred as a Gabor synthesis matrix.

Basics of Compressed Sensing

- A vector $x \in \mathbb{R}^M$ is $k-$**sparse** if it has k nonzero coordinates. That is, $\|x\|_0 := |\{i \mid x_i \neq 0\}| = k < M$
- One of the central problems in CS is that of reconstructing an unknown sparse vector $x \in \mathbb{R}^M$ from the linear measurements $y' = (\langle x, \psi_1 \rangle, \ldots, \langle x, \psi_m \rangle) \in \mathbb{R}^m$
- One can recover sparse x from its linear measurements by solving the following optimization problem:

$$P_0 : \min_x \|x\|_0 \text{ subject to } \quad \Psi x = y \tag{1}$$

- D.Donoho et.al. posed an equivalent of this problem as

$$P_1 : \min_x \|x\|_1 \text{ subject to } \quad y = \Psi x \tag{2}$$

Problem statement

- The authors posed the following three equivalent problems
- Determine the coefficient sequence of a vector

$$y = \sum_{\lambda \in \mathbb{Z}_n \times \mathbb{Z}_n} x_\lambda \Pi(\lambda) g,$$

  that is known to have a sparse representation in the Gabor system $\{\Pi(\lambda)g; \lambda \in \mathbb{Z}_n \times \mathbb{Z}_n\}$ with window $g \neq 0$.
- Establish the applicability of Ψ_g as measurement matrix for compressed sensing, that is, they aim at efficiently recovering an unknown signal x that is sparse in the Euclidean basis from its measurements $y = \Psi_g x$

Problem statement

- Identify from a single input output pair $(g, \Gamma g)$ the coefficient vector x of an operator

$$\Gamma = \sum_{\lambda \in \mathbb{Z}_n \times \mathbb{Z}_n} x_\lambda \Pi(\lambda),$$

where Γ is assumed to have a sparse representation in the system of time frequency shift matrices $\{\Pi(\lambda); \lambda \in \mathbb{Z}_n \times \mathbb{Z}_n\}$. That is, identify $\Gamma \in \mathbb{C}^{n \times n}$, or equivalently x, from its action $y = \Gamma g$ on a single vector g.

Advantages with matrix Ψ_g

- In the main results, the window vector $g \in \mathbb{C}^n$ is chosen at random, which implies that the measurement matrix Ψ_g depends only on n independent random variables as compared to $n \times N$ independent random variables in the case of Gaussian or Bernoulli measurement matrices.

Advantages with matrix Ψ_g

- The structure of Ψ_g allows for fast Fourier transform based matrix vector multiplication algorithms. This leads to efficient implementations of l_1-minimization methods

The window vector g

- The entries of Alltop window g^A can be defined as follows:

$$g_q^A = \frac{1}{\sqrt{n}} e^{\frac{2\pi i q^3}{n}}, q = 0, 1, \ldots, n-1.$$

- Through out the paper, the authors considered the randomly generated window g^R with entries

$$g_q^R = \frac{1}{\sqrt{n}} \epsilon_q, q = 0, 1, \ldots, n-1,$$

where the ϵ_q are independent and uniformly distributed on the torus $\{z \in \mathbb{C}, |z| = 1\}$.

Existing results

- In their earlier work[a], the authors had given the theoretical guarantees for basis pursuit algorithm when the sensing matrix is Gabor synthesis
- First, they have given the theoretical guarantees using coherence based arguments
- Since, coherence based arguments are worst case analysis, they gone through the average case analysis and provided the following recovery guarantees for BP algorithm
- BP can recover the $s-$sparse vector with high probability provided

$$s \leq C \frac{n}{\log(n)^u},$$

for some constant C, where $u = 1$ in the case of g^A and $u = 2$ in the case of g^R

[a] Pfander, G.E., Rauhut, H., Tanner, J.: Identification of matrices having a sparse representation. IEEE Trans. Signal Process. 56(11), 5376-5388 (2008)

Main results

- In the above result, the average case analysis means, the support set of x and and the signs of its nonzero coefficients are chosen at random
- In the present paper, the authors worked with the randomly generated window g^R and improved the above condition to $s \leq C\frac{n}{\log(n)}$ for deterministic x.

Theorem

Let n, be even and let $\Lambda \subset \mathbb{Z}_n \times \mathbb{Z}_n$ be of cardinality $|\Lambda| = s$. Let x with $supp(x) = \Lambda$ be such that on Λ the random phases $(sgn(x_\lambda))_{\lambda\in\Lambda}$ are independent and uniformly distributed on the torus $\{z \in \mathbb{C}, |z| = 1\}$. Let $\sigma > 8$. Choose the window $g = g^R$, with random entries independently and uniformly distributed on the tours. Then with probability at most

$$2(n^2 - s)exp\left(-\frac{n}{8\sigma s\log(n)}\right) + Csexp\left(-\frac{n}{16es}\right) + 4n^{-(\frac{\sigma}{4}-2)}$$

Main results

Basis pursuit algorithm fails to recover x from $y = \Psi_g x$. Here support of x is deterministic and the phase of the coefficient vector are random.

- The restriction to n even was made for the sake of simple exposition; a similar result holds also for n odd.

Theorem

Assume x is an arbitrary $s-$sparse coefficient vector. Choose the random unimodular Gabor window $g = g^R$ as stated above. Assume that

$$s \leq C \frac{n}{\log(\frac{n}{\epsilon})}$$

for some constant C. Then with probability at least $1 - \epsilon$ BP recovers x from $y = \Psi_g x$.

Stirling number of the first kind

Definition

The associated Stirling number of the first kind, denoted by $d_2(m, s)$ is the number of permutations of m elements which involve exactly s disjoint cycles and where each cycle has at least 2 elements.

- $d_2(0, 0) = 0, d_2(m, 0) = 0, d_2(m, s) = 0, m \geq 1, s > \frac{m}{2}$
- $d_2(m+1, s) = m[d_2(m, s) + d_2(m-1, s-1)], 1 \leq s \leq \frac{m}{2}$
- $d_2(m+1, s) \leq (2m)^{m-s}$.

Well conditioned sub matrices of Gabor synthesis matrices

- Many results on sparse recovery rely on the fact that small column submatrices of measurement or synthesis matrices such as Ψ_g are well-conditioned.

Theorem

Let $\epsilon, \delta \in (0,1)$ *and* $\Lambda = s$. *Suppose that*

$$s \leq \frac{\delta^2 n}{4e(\log \frac{s}{\epsilon}) + c}$$

with $c = log(\frac{e^2}{4(e-1)})$. *Then* $\|I_\Lambda - \Psi_\Lambda^* \Psi_\Lambda\| \leq \delta$ *with probability atleast* $1 - \epsilon$; *in other words the minimal and maximal eigenvalues of* $\Psi_\Lambda^* \Psi_\Lambda$ *satisfy* $1 - \delta \leq \lambda_{min} \leq \lambda_{man} \leq 1 + \delta$ *with probability at least* $1 - \epsilon$.

Proof.

Set $H_\Lambda = \Psi_\Lambda^* \Psi_\Lambda - I_\Lambda$. We want to show $\mathbb{P}(\|I_\Lambda - \Psi_\Lambda^* \Psi_\Lambda\| > \delta) \leq \epsilon$.
$\mathbb{P}(\|I_\Lambda - \Psi_\Lambda^* \Psi_\Lambda\| > \delta) = \mathbb{P}(\|H_\Lambda\| > \delta) = \mathbb{P}(\|H_\Lambda\|^{2m} > \delta^{2m})$
$\leq \delta^{-2m}\mathbb{E}[\|H_\Lambda\|^{2m}] = \delta^{-2m}\mathbb{E}[\|H_\Lambda^m\|^2] \leq \delta^{-2m}\mathbb{E}[\|H_\Lambda^m\|_F^2]$
$= \leq \delta^{-2m}\mathbb{E}[TrH_\Lambda^{2m}]$. □

Lemma

If $s = |\Lambda|$, $H_\Lambda = \Psi_\Lambda^* \Psi_\Lambda - I_\Lambda$ and m even then

$$\mathbb{E}[TrH_\Lambda^m] \leq s\left(\frac{s}{n}\right)^m \sum_{s=1}^{\frac{m}{2}} d_2(m,s)\left(\frac{n}{s}\right)^s.$$

Let $\psi_\lambda = \Pi(\lambda)g$ be the column of Ψ indexed by λ. By $R_\Lambda x$ we denote the restriction of a vector to the index set Λ.

Lemma

Suppose that $y = \Psi x$ for some x with $supp(x) = \Lambda$. If

$$|\langle \Psi_\Lambda^\dagger \psi_\rho, R_\Lambda sgn(x)\rangle| < 1, \forall \rho \notin \Lambda,$$

then x is the unique solution for the BP algorithm.

Proof of the main theorem 1

By using above lemma they calculated the failure probability of recovery. That is, they find the $\mathbb{P}(|\langle \Psi_\Lambda^\dagger \psi_\rho, R_\Lambda sgn(x)\rangle| \geq 1 \text{ for some} \rho \notin \Lambda)$ by using the following Bernstein type inequality for a sequence of independent random variables ϵ_k having uniform distribution on the torus,

$$\mathbb{P}\left(\left|\sum_j \epsilon_j a_j\right| \geq u\|a\|_2\right) \leq \frac{e^{-\kappa u^2}}{1-\kappa}$$

Thank you