EE 381V: Sparsity, Structures and Algorithms	**Spring 2010**

Lecture 20 — March 31

Lecturer: Sujay Sanghavi *Scribe: vishvAs vAsuki, Praneeth Netrapalli*

20.1 Topics covered

- Sparse Representation in Over Complete Basis
- Finding Sparse Representation using ℓ_1 minimization

20.2 Sparse representation in overcomplete basis

20.2.1 Motivation from digital signal processing

Let Φ be a matrix whose columns are composed of the basis functions/ vectors for the frequency domain. Let Ψ be a matrix similarly formed using the basis functions/ vectors for the time domain.

Consider the signal shown in figure 20.1. We have added together two signals with sparse representations in Φ and Ψ alone to produce a signal which is sparse in neither. In other words, this resultant signal does not have a sparse representation in Φ alone or in Ψ alone. But, it can be seen to have a sparse representation in Φ and Ψ together.

It must be noted, however that the columns of $[\Phi\ \Psi]$ are not linearly independent. So, $[\Phi\ \Psi]$ forms an 'overcomplete basis'. This motivates the following problem.

20.2.2 Problem

Given a signal (or a vector x in general), when can we find a sparse representation in the 'overcomplete basis' $[\Phi\ \Psi]$?

Ill posedness/ non-uniqueness

There are an infinite number of representations for any signal in a pair of bases. Looking for sparsity in representation may help, but this is not enough to guarantee uniqueness.

To see this, consider a basis Ψ which differs from Φ in only two basis vectors, but is the same as Φ otherwise. Then, sparsity of x in Ψ is equivalent to sparsity in Φ; and it is even hard for an x which is sparse in Ψ or Φ alone, to have a unique representation in the overcomplete basis $[\Phi\ \Psi]$. For example, suppose that both Φ and Ψ include the basis vectors $\{\phi_1, \phi_2\}$, and consider $x = \phi_1 + \phi_2$. x has an infinite number of sparse representations in the basis $[\Phi\ \Psi]$.

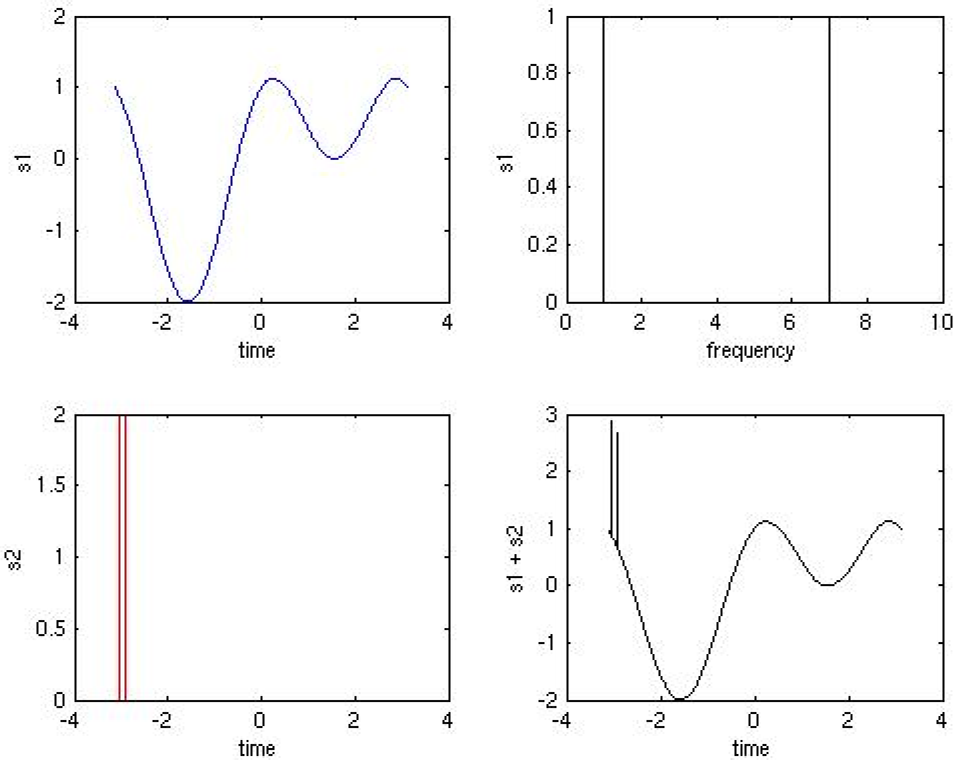

Figure 20.1. The signal s1 has a sparse representation in the frequency domain. s2 has a sparse representation in the time domain. But, the signal (s1 + s2) has a sparse repersentation in an overcompelte basis, but not in either the frequency or time domain alone.

So, we must add an additional condition: Φ and Ψ must be 'different' enough. This property is usually referred to as 'incoherence'.

20.2.3 Basic uncertainty principle

Theorem 20.1. *Let Φ and Ψ be two orthonormal bases for a vector space. For any vector $s \neq 0$, let $s = \Phi a$ and $s = \Psi b$. Let $M = \max_{i,j} |\langle \phi_i, \psi_j \rangle|$. Then:*

$$\frac{\|a\|_0 + \|b\|_0}{2} \geq \sqrt{\|a\|_0 \|b\|_0} \geq M^{-1}$$

.

Proof: The first inequality comes from the well known AM-GM inequality, so we prove $\sqrt{\|a\|_0 \|b\|_0} \geq M^{-1}$.

Without loss of generality, suppose that $s^T s = 1$.

$$1 = a^T \Phi^T \Psi b \leq \sum_{i,j} |a_i||\langle \phi_i, \psi_j \rangle||b_j| \leq M \sum_i |a_i| \sum_j |b_j|$$

Let $A = \|a\|_0, B = \|b\|_0$. From $s^T s = 1$, we have that $\sum_i |a_i|^2 = 1 = \sum_i |b_i|^2$. Also, $\forall i : |a_i| \leq 1, |b_i| \leq 1$. So:

$$\sum_i |a_i| \leq \sqrt{A}, \sum_i |b_i| \leq \sqrt{B}$$

Substituting this into the earlier inequality, we have $\sqrt{\|a\|_0 \|b\|_0} \geq M^{-1}$. □

From this, we have the following theorem, whose proof is left as a homework exercise.

Theorem 20.2. *If $s = [\Phi\ \Psi]g_1, s = [\Phi\ \Psi]g_2$, then $\|g_1\|_0 + \|g_2\|_0 \geq \frac{2}{M}$.*

From this, we have the following corollary:

Corollary 20.3. *Suppose that $s = [\Phi\ \Psi]g_1$, and $\|g_1\|_0 < M^{-1}$. Then, there does not exist a g_2 such that $s = [\Phi\ \Psi]g_2$ and $\|g_2\|_0 < M^{-1}$.*

Incoherence is the situation where M is small. Thus, incoherence ensures uniqueness for the sparsest representation. In other words, any other representation will have more non-zeros.

This motivates the following question: If s has a unique sparse representation, how do we find it? The answer to this is explored in the next section.

20.3 Sparse Representation using ℓ_1 minimization

Suppose we want to find the sparsest γ such that $[\Phi\Psi]\,\gamma = s$. The set of all feasible γ is an affine subspace of $\mathbb{R}^n$. We want to know when ℓ_1 minimization gives us the sparsest solution.

Theorem 20.4 (Donoho and Huo). *If $\|\gamma\|_0 < \frac{1}{2}(1+\frac{1}{M})$ then $\|\widetilde{\gamma}\|_1 > \|\gamma\|_1$ for all $\widetilde{\gamma}$ such that $[\Phi\Psi]\,\widetilde{\gamma} = [\Phi\Psi]\,\gamma$*

Proof: Let $\|\gamma\|_0 < \frac{1}{2}(1 + \frac{1}{M})$. Consider any $\widetilde{\gamma}$ such that $[\Phi\Psi]\,\widetilde{\gamma} = [\Phi\Psi]\,\gamma$. Let $x = \widetilde{\gamma} - \gamma$. Then, we have $[\Phi\Psi]\,x = 0$.

$$\|\widetilde{\gamma}\|_1 > \|\gamma\|_1$$

$$\Leftrightarrow \sum_{k \in \text{ supp } (\gamma)} (|\gamma_k + x_k| - |\gamma_k|) + \sum_{k \notin \text{ supp } (\gamma)} |x_k| > 0$$

$$\Leftarrow \sum_{k \in \text{ supp } (\gamma)} -|x_k| + \sum_{k \notin \text{ supp } (\gamma)} |x_k| > 0$$

$$\Leftrightarrow \frac{\sum_{\text{ON}} |x_k|}{\sum_{\text{ALL}} |x_k|} < \frac{1}{2}$$

where $\sum_{\text{ON}} |x_k| = \sum_{k \in \text{ supp } (\gamma)} |x_k|$ and $\sum_{\text{ALL}} |x_k| = \sum_{k} |x_k|$.

We want this to be true for all supports of size less than $\frac{1}{2}(1 + \frac{1}{M})$ and for all x such that $[\Phi\Psi]\, x = 0$. Let $x \neq 0$ be any such vector and let i be such that $|x_i| \geq |x_j|$ for all j. WLOG, by rescaling, we can assume $x_i = v$ where v is a fixed number. For this, we have $\sum_{\text{ON}} |x_k| \leq \|x\|_0 |v|$.

$$x = \begin{bmatrix} x^{\Phi} \\ x^{\Psi} \end{bmatrix}$$
$$\|x\|_1 = \|x^{\Phi}\|_1 + \|x^{\Psi}\|_1$$

WLOG, say i is in the Φ part.

$$\Phi x^{\Phi} = -\Psi x^{\Psi}$$
$$x^{\Phi} = -\Phi^T \Psi x^{\Psi}$$

since Φ is an orthonormal matrix. So we have,

$$\begin{aligned} |v| = |x_i| &= | \left[\Phi^T \Psi\right]_i x^{\Psi}| \\ &\leq M \|x^{\Psi}\|_1 \end{aligned}$$

$$\Rightarrow \|x^{\Psi}\|_1 \geq \frac{|v|}{M}$$

We also have,

$$\begin{aligned} \|x^{\Phi}\|_1 &\geq |v| \\ \Rightarrow \sum_{\text{ALL}} |x_k| &= \|x\|_1 \geq |v| \left(1 + \tfrac{1}{M}\right) \end{aligned}$$

So for any x such that $[\Phi\Psi]\, x = 0$,

$$\frac{\sum_{\text{ON}} |x_k|}{\sum_{\text{ALL}} |x_k|} \leq \frac{\|\gamma\|_0}{1 + \frac{1}{M}}$$

and since $\|\gamma\|_0 < \frac{1}{2}\left(1 + \frac{1}{M}\right)$, we have

$$\frac{\sum_{\text{ON}} |x_k|}{\sum_{\text{ALL}} |x_k|} < \frac{1}{2}$$

proving the result.

$\square$

Elad and Bruckstein further refine the sufficient condition to $\|\gamma\|_0 < \frac{\sqrt{2}-\frac{1}{2}}{M}$. So under these conditions, if we solve argmin $\|\gamma\|_1$ s.t. $[\Phi\Psi]\,\gamma = s$, we are guaranteed to find the unique sparsest solution.