Harmonic Analysis: Theory and Applications in Advanced Signal Processing
Spring semester 2015

Homework 6 - Solutions available on June 2, 2015

Problem 1  Recovery of Approximately Sparse Signals Using \(\ell_1\)-Minimization

Let \(D \in \mathbb{C}^{K \times N}\) be a matrix \((K < N)\). For an index set \(\Lambda \subseteq [1, N]\), define the quantity

\[
C(D, \Lambda) = \max_{\substack{h \in \mathbb{N}(D) \setminus \{0\}}} \frac{\|h_\Lambda\|_1}{\|h\|_1} = \max_{\substack{h \in \mathbb{N}(D) \setminus \{0\}}} \frac{\sum_{k \in \Lambda} |h_k|}{\sum_{k=1}^N |h_k|},
\]

where \(h_\Lambda\) is the vector constructed from \(h\) by setting to zero all but the entries indexed by \(\Lambda\).

Let \(x \in \mathbb{C}^N\) and consider the \(\ell_1\)-minimization problem:

\[
(P1) \quad \text{minimize } \|\tilde{x}\|_1 \quad \text{subject to } \quad Dx = Dx.
\]

In the lecture, we have seen that if \(x\) is \(s\)-sparse with support \(S = \text{supp}(x) = \{k \in [1, N] : x_k \neq 0\}\), \(|S| \leq s\), and if \(C(D, S) < 1/2\), then \(x\) is the unique solution to (P1).

Now, assume that \(x\) is approximately sparse, meaning that many of the entries of \(x\) are close to zero. Show that if \(C(D, S) < 1/2\), the solution \(x^*\) to (P1) satisfies

\[
\|x^* - x\|_1 \leq \frac{2}{1 - 2C(D, S)} \|x - x_S\|_1,
\]

where \(S\) denotes the set containing the indices of the \(s\) largest (in magnitude) components of \(x\), the vector \(x_S\) thus representing the best \(s\)-sparse approximation to \(x\). Do you recover the condition derived in the lecture for the case of exactly sparse signals?

Problem 2  Restricted Isometry Property and Coherence

Let \(A \in \mathbb{C}^{M \times N}\) be a matrix having normalized columns, i.e., each column \(a_\ell, \ell \in [1, N]\), satisfies \(\|a_\ell\|_2 = 1\). For \(s \in [1, N]\), we say that \(A\) satisfies the restricted isometry property (RIP) of order \(s\) if there exists \(\delta \in (0, 1)\) such that

\[
(1 - \delta) \|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta) \|x\|_2^2
\]

for all \(x \in \mathbb{C}^N\) such that \(|\text{supp}(x)| \leq s\) and we define \(\delta_s(A)\) to be the smallest such \(\delta\).

a) Show that if \(A\) satisfies the RIP of order \(s\), then it also satisfies the RIP of order \(s' \leq s\).
b) Show that
\[ \delta_s(A) = \max_{S \subseteq [1,N], |S| \leq s} \left\| A_S^H A_S - I_s \right\|_{2 \to 2}, \]
where \( A_S \) denotes the matrix consisting of the columns of \( A \) indexed by the set \( S \subseteq [1,N] \) and \( A_S^H \) denotes the conjugate transpose of \( A_S \).

Recall: The matrix norm \( \| \cdot \|_{2 \to 2} \) is defined as
\[ \| A \|_{2 \to 2} = \sup_{x \in \mathbb{C}^N, x \neq 0} \frac{\| A x \|_2}{\| x \|_2}, \]
for \( A \in \mathbb{C}^{N \times N} \).

c) What does \( \delta_s(A) \) say about the eigenvalues of the matrix \( A_S^H A_S \), where \( S \) is a subset of \([1,N]\) of size \( s \)? Find an upper bound for the condition number of \( A_S^H A_S \) expressed in terms of \( \delta_s(A) \).

d) Prove that
\[ \delta_s(A) \leq \mu(A)(s - 1), \]
where
\[ \mu(A) = \max_{k,\ell \in [1,N], k \neq \ell} |\langle a_k, a_\ell \rangle| \]
is the coherence of the matrix \( A \).

**Problem 3** Restricted Isometry Property: A Counterexample

Let \( A \in \mathbb{C}^{M \times N} \) be the matrix with constant elements
\[ a_{k,\ell} = \frac{1}{\sqrt{M}}, \quad k \in [1,M], \ell \in [1,N]. \]
Show that \( A \) does not satisfy the RIP of order \( s \geq 2 \).

**Problem 4** Orthogonal Matching Pursuit

Let \( A \in \mathbb{R}^{M \times N} \) be a measurement matrix and \( x \in \mathbb{R}^N \) an \( s \)-sparse vector, i.e., \( x \) has at most \( s \) nonzero entries. Define \( y = Ax \in \mathbb{R}^M \).

Implement in Matlab the orthogonal matching pursuit (OMP) algorithm described in Discussion Session 11, which enables to recover the vector \( x \) from the measurement vector \( y \), knowing the measurement matrix \( A \) and the sparsity level \( s \) only.

For concreteness, take \( N = 512 \) and \( M = 128 \). Generate a random sparse vector \( x \in \mathbb{R}^N \) with a sparsity level \( s = 20 \) and a random measurement matrix \( A \in \mathbb{R}^{M \times N} \) with entries identically and independently distributed according to a Gaussian distribution of zero mean and variance \( 1/M \). Compute \( y = Ax \) and reconstruct \( x \) from \( A \) and \( y \) using your OMP algorithm. Repeat the procedure with other values of \( s \).
Recall: The key to OMP is to determine which columns of $A$ participate in the measurement vector $y$. The idea behind this is to pick columns of $A$ in a greedy fashion. At each iteration, we choose the column of $A$ which is most correlated with the residual, i.e., the remaining part of $y$ which has not yet been approximated. This contribution is then subtracted from $y$ and the algorithm iterates on the residual. This method is summarized in Algorithm 1.

**Algorithm 1: Orthogonal Matching Pursuit (OMP)**

**Input**: $A$, a measurement matrix in $\mathbb{R}^{M \times N}$, $y$, a vector of measurements in $\mathbb{R}^{M}$, $s$, the sparsity level of the ideal signal.

**Output**: $\hat{x}$, a sparse estimate of the ideal signal in $\mathbb{R}^{N}$, $I$, the support of the estimated signal, i.e., the set containing the position of the nonzero elements of $\hat{x}$.

1: Initialize the signal estimate $x^{(0)} = 0$, the index set $I^{(0)} = \emptyset$, the matrix of chosen atoms $A^{(0)} = []$, and the iteration counter $t = 1$.
2: while $t < s$ do
3: Calculate the residual:
   \[ r^{(t)} = y - A^{(t-1)}x^{(t-1)}. \]
4: Find the index of the column of $A$ that is most correlated with $r^{(t)}$:
   \[ i^{(t)} = \arg\max_{j=1,\ldots,N} |\langle r^{(t)}, a_j \rangle|. \]
   If the maximum occurs for multiple indices, choose one arbitrarily.
5: Augment the index set $I^{(t)} = I^{(t-1)} \cup \{i^{(t)}\}$ and the matrix of chosen atoms $A^{(t)} = \text{concat}(A^{(t-1)}, a_{i^{(t)}})$.
6: Update the signal estimate by solving the least squares problem:
   \[ x^{(t)} = \arg\min_{x \in \mathbb{R}^{N}} \|y - A^{(t)}x\|_2, \]
   i.e., $x^{(t)} = A^{(t)}\dagger y$.
7: $t = t + 1$.
8: end while
9: $\hat{x} = x^{(t)}$ and $I = I^{(t)}$.
10: return $\hat{x}, I$.