Homework 7

Assigned: 08/11/2011.
Due: 18/11/2011.

Exercise 1. ℓ_1-MINIMIZATION APPROXIMATION GUARANTEES
Let $u = \Phi x^* + n$ be a set of observations where $u \in \mathbb{R}^m$ and $\Phi \in \mathbb{R}^{m \times n}$, $(m < n)$. Here, $x^* \in \mathbb{R}^n$ is the signal of interest: it can be exactly k-sparse ($k < m$) or compressible, i.e. we can well-approximate x^* by keeping the k largest (in magnitude) coefficients and (yet) have a good approximation of x^*. We denote the best k-sparse approximation of x^* as x^k.
One way to efficiently reconstruct x^* is by using ℓ_1-minimization schemes, according to which we try to solve the following optimization problem:

$$\min_{x: x \in \mathbb{R}^n} \|x\|_1 \quad \text{subject to} \quad \|u - \Phi x\|_2 \leq \varepsilon. \tag{1}$$

We reserve $\tilde{x} \in \mathbb{R}^n$ to denote the sparse solution of (1). In this exercise, we prove the following approximation guarantees for the ℓ_1-minimization problem:

Theorem 1 *Assume $\Phi \in \mathbb{R}^{m \times n}$ satisfies the RIP for $\delta_{2k} < \sqrt{2} - 1$ and let $\|n\|_2 \leq \varepsilon$. Given positive constants C_1, C_2, the solution $\tilde{x}$ satisfies the following approximation guarantees:*

$$\|\tilde{x} - x^*\|_2 \leq \frac{C_1}{\sqrt{k}} \|x^k - x^*\|_1 + C_2 \varepsilon. \tag{2}$$

To prove Theorem 1, we prove a series of steps that lead to the desired result. An important "ingredient" in the proof is described next:

Lemma 1 *For all $x_1, x_2 \in \mathbb{R}^n$ with disjoint support sets $T_1, T_2 \subseteq \mathcal{N}$ (respectively) such that $|T_1| \leq k_1$ and $|T_2| \leq k_2$, the following inequality holds:*

$$|\langle \Phi x_1, \Phi x_2 \rangle| \leq \delta_{k_1 + k_2} \|x_1\|_2 \|x_2\|_2. \tag{3}$$

To set-up notation, the index set of n-dimensional vectors is denoted as $\mathcal{N} = \{1, 2, \ldots, n\}$. For $\mathcal{S} \subseteq \mathcal{N}$, we define the complement set $\mathcal{S}^c = \mathcal{N} \setminus \mathcal{S}$. Moreover, given a set $\mathcal{S} \subseteq \mathcal{N}$ and a vector $x \in \mathbb{R}^n$, $(x)_{\mathcal{S}} \in \mathbb{R}^n$ denotes a vector with nonzero coefficients at the positions indexed by $\mathcal{S}$.

1. **[3 points]** Let $\alpha := \tilde{x} - x^* \in \mathbb{R}^n$. We define the following index sets: $\Gamma_i \subseteq \mathcal{N}$, $i = 1, 2, \ldots$, where Γ_0 contains the indices of the k largest (in magnitude) elements of x^*, Γ_1 contains the indices of the *next* k largest elements in α *outside* Γ_0, Γ_2 contains the indices of the *next* k largest elements in α *outside* $\Gamma_0 \cup \Gamma_1$ etc. Thus,

   $$\Gamma_i \cap \Gamma_j = \{\emptyset\}, \ \forall i \neq j \ \text{ and } \ \alpha = \sum_i (\alpha)_{\Gamma_i}.$$

   Given the above:

   - Justify the following set of inequalities for $i \geq 2$:

   $$\|(\alpha)_{\Gamma_i}\|_2 \leq \sqrt{k} \|(\alpha)_{\Gamma_i}\|_\infty \leq \frac{\|(\alpha)_{\Gamma_{i-1}}\|_1}{\sqrt{k}}, \tag{4}$$

and prove:

$$\|(\alpha)_{(\Gamma_0\cup\Gamma_1)^c}\|_2 \leq \frac{\|(\alpha)_{\Gamma_0^c}\|_1}{\sqrt{k}}. \tag{5}$$

Hint: Find the connection between $\|(\alpha)_{(\Gamma_0\cup\Gamma_1)^c}\|_2$ and $\sum_{i\geq 2}\|(\alpha)_{\Gamma_i}\|_2$ and then use (4) to prove (5).

2. **[1 point]** Given that $\tilde{x}$ is a feasible solution, i.e., satisfies (1), prove:

$$\|\Phi\tilde{x} - \Phi x^*\|_2 \leq 2\varepsilon. \tag{6}$$

3. Since:

$$\|\alpha\|_2 \leq \|(\alpha)_{\Gamma_0\cup\Gamma_1}\|_2 + \|(\alpha)_{(\Gamma_0\cup\Gamma_1)^c}\|_2, \tag{7}$$

we want to bound the quantities $\|(\alpha)_{\Gamma_0\cup\Gamma_1}\|_2$ and $\|(\alpha)_{(\Gamma_0\cup\Gamma_1)^c}\|_2$.

- **[2 points]** Prove:

$$\|(\alpha)_{(\Gamma_0\cup\Gamma_1)^c}\|_2 \leq \|(\alpha)_{\Gamma_0}\|_2 + \frac{2}{\sqrt{k}}\|x^k - x^*\|_1. \tag{8}$$

Hint: Use the fact that $\|x^*\|_1 \geq \|\tilde{x}\|_1$ and inequality (5).

- **[3 points]** Prove:

$$\|(\alpha)_{\Gamma_0\cup\Gamma_1}\|_2 \leq \frac{2\varepsilon\sqrt{1+\delta_{2k}}}{1-\delta_{2k}} + \frac{\sqrt{2}\delta_{2k}}{1-\delta_{2k}}\frac{\|(\alpha)_{\Gamma_0^c}\|_1}{\sqrt{k}}, \tag{9}$$

using the following steps:

(a) Using $\Phi(\alpha)_{\Gamma_0\cup\Gamma_1} = \Phi\alpha - \sum_{i\geq 2}\Phi(\alpha)_{\Gamma_i}$, prove:

$$\|\Phi(\alpha)_{\Gamma_0\cup\Gamma_1}\|_2^2 = \langle\Phi(\alpha)_{\Gamma_0\cup\Gamma_1}, \Phi\alpha\rangle - \langle\Phi(\alpha)_{\Gamma_0\cup\Gamma_1}, \sum_{i\geq 2}\Phi(\alpha)_{\Gamma_i}\rangle. \tag{10}$$

(b) Prove $|\langle\Phi(\alpha)_{\Gamma_0\cup\Gamma_1}, \Phi\alpha\rangle| \leq 2\varepsilon\sqrt{1+\delta_{2k}}\|(\alpha)_{\Gamma_0\cup\Gamma_1}\|_2$.

(c) Using Lemma 1, verify that $|\langle\Phi(\alpha)_{\Gamma_i}, \Phi(\alpha)_{\Gamma_j}\rangle| \leq \delta_{2k}\|(\alpha)_{\Gamma_i}\|_2\|(\alpha)_{\Gamma_j}\|_2$.

(d) Given that $\|(\alpha)_{\Gamma_0}\|_2 + \|(\alpha)_{\Gamma_1}\|_2 \leq \sqrt{2}\|(\alpha)_{\Gamma_0\cup\Gamma_1}\|_2$, prove (9).
Hint: Use the definition of RIP on $(\alpha)_{\Gamma_0\cup\Gamma_1}$.

4. **[1 point]** Given the derivations above, complete the proof of Theorem 1 using (7) and give the restricted isometry constant constraints in terms of δ_{2k}.

Exercise 2. Proving RIP for Random matrices
In this exercise we will provide a simple proof of a fundamental CS construct namely the Restricted Isometry property (RIP). Let Φ denote a $n \times N$ matrix where $n < N$. We say that Φ satisfies the RIP of order k if there exists $\delta_k \in (0,1)$ such that

$$(1-\delta_k)\|\mathbf{x}_T\|_2^2 \leq \|\Phi_T\mathbf{x}_T\|_2^2 \leq (1+\delta_k)\|\mathbf{x}_T\|_2^2$$

holds *for all* sets $T \subset \{1,\ldots,N\}$ with $|T| \leq k$ where T denotes the set of column indices. Φ_T is the $n \times |T|$ matrix composed of those columns. $\mathbf{x}_T$ denotes the vector obtained by retaining only the entries of $\mathbf{x}$ corresponding to the column indices T. Denote by X_T the set of all vectors in $\mathbb{R}^N$ that are zero outside of T. Assume that the matrix Φ has i.i.d entries with

$$\mathbb{E}[\phi_{i,j}] = 0 \quad \text{and} \quad \mathbb{E}[\phi_{i,j}^2] = \frac{1}{n}.$$

1. **[2 points]** Show that $\mathbb{E}[\|\Phi\mathbf{x}\|_2^2] = \|\mathbf{x}\|_2^2$ for any given $\mathbf{x} \in \mathbb{R}^N$.

Now for any $\mathbf{x} \in \mathbb{R}^N$ the random variable $\|\Phi\mathbf{x}\|_2^2$ is strongly concentrated around its mean i.e. for any $0 < \epsilon < 1$ we have

$$\mathbb{P}\left(\left|\|\Phi\mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2\right| \geq \epsilon \|\mathbf{x}\|_2^2\right) \leq 2e^{-nc_0(\epsilon)} \tag{11}$$

where $c_0(\epsilon)$ is a constant depending only on ϵ and such that $c_0(\epsilon) > 0$ for all $\epsilon \in (0,1)$. We make use of this inequality in the following parts.

2. **[4 points]** As Φ is linear we can assume that $\|\mathbf{x}\|_2 = 1$ for all $\mathbf{x} \in X_T$. Assume $|T| = k$. Now it is known from covering numbers that there exists $Q_T \subseteq X_T$ with $|Q_T| \leq (\frac{12}{\delta})^k$ where $\delta \in (0,1)$ and $\|q\|_2 = 1$ for all $q \in Q_T$ so that

$$\min_{q \in Q_T} \|\mathbf{x} - \mathbf{q}\|_2 \leq \frac{\delta}{4} \tag{12}$$

holds for any $\mathbf{x} \in X_T$. Use (12) along with the concentration result in (11) to show that *for any given* T with $|T| = k$,

$$(1-\delta)\|\mathbf{x}\|_2 \leq \|\Phi\mathbf{x}\|_2 \leq (1+\delta)\|\mathbf{x}\|_2$$

holds $\forall \mathbf{x} \in X_T$ with probability at least $1 - 2\left(\frac{12}{\delta}\right)^k e^{-c_0(\delta/2)n}$.

3. **[4 points]** Use the result from previous part to show that for a given $0 < \delta < 1$, n, N, there exist constants $c_1, c_2 > 0$ depending only on δ such that *for all* sets T with $|T| = k$,

$$(1-\delta)\|\mathbf{x}_T\|_2 \leq \|\Phi_T\mathbf{x}_T\|_2 \leq (1+\delta)\|\mathbf{x}_T\|_2$$

holds for the prescribed δ and any $k \leq \frac{c_1 n}{\log(N/k)}$ with probability at least $1 - 2e^{-c_2 n}$. Conclude that Φ satisfies the RIP of order k with high probability. (Hint: Use $\binom{N}{k} \leq (\frac{eN}{k})^k$.)