CO 769: Topics in Continuous Optimization
Recent Advances in Compressive Sensing, ℓ_1 minimization, and Rank Minimization
Winter 2009

PROBLEM SET 2

S. Vavasis

Handed out: 2009-Mar-5
Due: 2009-April-8
Please answer at least six of these questions.

1. Download the lecture notes entitled "Elementary proof of the spherical section property for random matrices" from the course webpage. Notice that the proof of Lemma 3 has been suppressed. Supply the missing proof.

   [Hint: Let γ stand for the largest absolute entry of $\mathbf{x}(\bar{T})$, and compute an inequality for $\|\mathbf{x}\|_1$ involving γ and Lemma 2.]

2. Same question for Lemma 16.

   [Hint: First prove it for the case $|T| = m'$ by a simple application of the union bound to Lemma 15. Then show that the $|T| = m'$ result implies $|T| > m'$ case (with a larger constant) by splitting an $m \times |T|$ matrix into $m \times m'$ blocks and using a result about $\sigma_{\max}(\cdot)$ of a matrix concatenation.]

3. The paper by Donoho and Stark ("Uncertainty Principles and Signal Recovery", *SIAM J. Appl. Math.* 49 (1989) 906–931) is sometimes cited as intuition for why compressive sensing works with partial Fourier information. Specifically, this paper proves that the discrete Fourier transform of a sparse signal is dense. This result suggests that a small sample of the entries of the Fourier transform may be enough to recover the signal.

   Find this paper on the web and read the paragraph at the bottom of p. 911: "Theorem 2 has an analogue for discrete time. The proof is the same, step-by-step..."

   State the theorem described by this paragraph and prove it.

4. In the Tropp and Gilbert paper on OMP, the authors need to solve an $m \times l$ linear least-squares problem on the lth iteration. The coefficient matrix is equal to the coefficient matrix from the preceding iteration with one additional column appended. They use Gram-Schmidt orthogonalization, and claim (p. 4657) that the incremental cost on the lth iteration for least-squares solution is $O(lm)$. Show that this time can be improved by using the Givens or Householder algorithms for least squares in an incremental fashion. [Look up these algorithms in a book on numerical linear algebra if you aren't familiar with them.]

5. The Tropp and Gilbert OMP algorithm can be classified as the weaker form of compressive sensing, meaning that the main theorem says that for every sparse signal, there is a high probability that most measurement matrices will work. [In contrast, strong compressive sensing would say that there is a single measurement matrix that works for all signals, and that a random matrix will work with high probability.] Explain why their proof works does not extend to the strong case, at least in an obvious way.

6. Develop linear-time algorithms for the following two problems. In both cases, demonstrate that your algorithm is correct by showing that the KKT conditions are satisfied at the optimizer.

   (a) $\min \mathbf{c}^T\mathbf{x}$ subject to $\|\mathbf{x}\|_1 \le 1$.

   (b) $\min \|\mathbf{x} - \mathbf{x}_0\|_2$ subject to $\|\mathbf{x}\|_1 \le 1$.

   These are subproblems that arise in first order proximal algorithms like Nesterov's.

7. The sparsest vector problem is as follows. Given an integer $m \times n$ matrix A and integer right-hand side m-vector $\mathbf{b}$, find $\mathbf{x}$ to minimize $\|\mathbf{x}\|_0$ subject to $A\mathbf{x} = \mathbf{b}$. Show that the sparsest vector problem is NP-hard. (This question and the next are due to Garey and Johnson.)

   [Hint: Use a reduction from three-dimensional matching (3DM). The 3DM problem was described in lecture and is one of the original NP-complete problems from Karp's paper, and its proof of NP-hardness is detailed in the Garey and Johnson book. You do not have to prove that 3DM is NP-complete.]

8. Show that the sparsest vector problem lies in NP. The nontrivial issue is bounding the number of bits in the certificate. For this purpose, a result due to Edmonds might be handy. Edmonds showed that if A is an $n \times n$ nonsingular integer matrix and $\mathbf{b}$ is an integer n-vector, then the number of bits needed to write down $A^{-1}\mathbf{b}$ (a rational vector, hence it may be written with fractions), is bounded by a polynomial in n, the number of bits in A and the number of bits in $\mathbf{b}$. Edmonds' proof is based on Cramer's rule.

9. In the Ames and Vavasis paper, rank minimization is defined to be the problem of minimizing the rank of a matrix X subject to linear inequality and equality constraints on the entries of X. Ames and Vavasis argue that the problem is NP-hard since clique can be reduced to it. Show that rank minimization is still NP-hard even if only linear equality constraints (no inequalities) are permitted.

   [Hint: there are several approaches; one simple approach is to note that the sparsest vector problem is a special case of rank minimization.]

10. Focus on the linear-equality constrained rank minimization problem. Assume all the data is integral. The rank minimization problem is not known to be in NP, but it is known to be decidable (i.e., a terminating Turing Machine algorithm exists for the problem). This is because the solution to rank minimization is also the solution to a system of (perhaps very many) polynomial equations and inequalities, and systems of polynomial equations and inequalities are known to be decidable in superexponential time. Explain how to reduce the decision version of rank minimization to a system of polynomial equations and inequalities. [Hint: use determinants.]