Limits of Approximation Algorithms 24 Apr, 2010

Problem Set 2

- Due Date: **13 May (Thurs), 2010**
- It is recommended that you try to solve all the exercises and problems, but you need to submit the writeup for only 5 of the 8 problems (note the length of the problem statement is not reflective of the difficulty of the problem!).
- Collaboration is encouraged, but all writeups must be done individually.
- Indicate names of all collaborators.
- Refering sources other than the lecture notes is discouraged, since for some of the problems a Google search will reveal the solution. But if you do use an outside source (text books, lecture notes, any material available online), do mention the same in your writeup.

Notation:

- $\mathbb{F}$ is a field of size q
- $\mathcal{S}_k^m$ is the set of affine subspaces of dimension k in $\mathbb{F}^m$.
- $P_{m,d}$ is the set of m-variate degree d polynomials

Exercises

1. **[Schwartz-Zippel]**

   If $p : \mathbb{F}^m \to \mathbb{F}$ is a non-zero m-variate polynomial of total degree at most d, show that

   $$\Pr_{x \in \mathbb{F}^m}[p(x) = 0] \leq \frac{d}{|\mathbb{F}|}.$$

2. **[Orthogonality via Schwartz-Zippel]**

   In class, we showed that $\mathbb{E}_{x\in\mathbb{F}^m}[\chi_\alpha(x)] = 0$ for $\alpha \neq (0, 0, \ldots, 0)$ where χ_α's are the characters defined as $\chi_\alpha(x_1, \ldots, x_m) = (-1)^{\sum \alpha_i x_i}$. Give an alternate proof using Schwartz-Zippel to the polynomial χ_α.

Problems

1. **[linearity test of 3 functions]**

   Consider the following modification of the BLR-linearity test towards testing linearity of 3 functions $f, g, h : \{0,1\}^n \to \{1,-1\}$ simultaneously.

   $$\mathsf{BLR\text{-}3\text{-}Test}^{f,g,h} : \text{“ } \begin{array}{l} \text{1. Choose } y, z \in_R \{0,1\}^n \text{ independently} \\ \text{2. Query } f(y), g(z), \text{ and } h(y+z) \\ \text{3. Accept if } f(y)g(z)h(y+z) = 1. \end{array} \quad \text{”}$$

   Clearly, if the three functions f, g, h are the same linear function, then the above test accepts with probability 1. Suppose one of the three functions f, g, h (say f) and its negation (i.e., $-f$) is δ-far from linear (this means $\max_\alpha |\hat{f}_\alpha| \le 1 - 2\delta$), show that

   $$\Pr_{y,z}[\mathsf{BLR\text{-}3\text{-}Test}^{f,g,h} \text{ rejects }] \ge \delta.$$

   [Hint: The Cauchy-Schwarz inequality $(\sum a_i b_i)^2 \le (\sum a_i^2) \cdot (\sum a_i^2)$ may come useful.]

2. **[recycling queries in linearity test]**

   In lecture, we analyzed the soundness of the $\mathsf{BLR\text{-}Test}$ to show that if f is $(1/2-\varepsilon)$-far from linear, then the test accepts with probability at most $1/2+\varepsilon$. If we repeat this test k times, we obtain a linearity test which makes $3k$ queries and has the following property: if f is $(1/2-\varepsilon)$-far from linear, then the test accepts with probability at most $(1/2+\varepsilon)^k = 1/2^k + \delta$. Thus every additional 3 queries improves the soundness by a factor of 1/2. In this problem, we show that this can be considerably improved.

   Assume that both f and $-f$ are $(1-\varepsilon)/2$-far from linear (i.e., $\max_\alpha |\hat{f}_\alpha| \le \varepsilon$). Consider the following linearity test (parameterized by k).

   $$\mathsf{Test}_k^f : \text{“ } \begin{array}{l} \text{1. Choose } z_1, z_2, \ldots, z_k \in_R \{0,1\}^n \\ \text{2. For each distinct pair } (i,j) \in \{1,\ldots,k\} \\ \qquad\qquad \text{Check if } f(z_i)f(z_j)f(z_i+z_j) = 1. \\ \text{3. Accept if all the tests pass.} \end{array} \quad \text{”}$$

   Observe that this test makes at most $k + \binom{k}{2}$ queries. We will show below that the soundness of the test is roughly $2^{-\binom{k}{2}}$, thus showing that every additional query improves the soundness by a factor of 1/2 (almost).

   Assume that both f and $-f$ are $(1-\varepsilon)/2$-far from linear.

(a) Show that the acceptance probability of the above test is given by

$$\begin{aligned}
\Pr[\mathsf{acc}] &= \mathbb{E}_{z_1,\ldots,z_k}\left[\prod_{i,j}\left(\frac{1+f(z_i)f(z_j)f(z_i+z_j)}{2}\right)\right] \\
&= \frac{1}{2^{\binom{k}{2}}}\cdot\sum_{S\subseteq\binom{[k]}{2}}\mathbb{E}_{z_1,\ldots,z_k}\left[\prod_{(i,j)\in S} f(z_i)f(z_j)f(z_i+z_j)\right]
\end{aligned}$$

(b) Consider any term in the above summation corresponding to a non-empty S (i.e., $\mathbb{E}_{z_1,\ldots,z_k}\left[\prod_{(i,j)\in S} f(z_i)f(z_j)f(z_i+z_j)\right]$). Suppose $(1,2)\in S$. Show that $\mathbb{E}_{z_1,\ldots,z_k}\left[\prod_{(i,j)\in S} f(z_i)f(z_j)f(z_i+z_j)\right]$ is upper bounded by $\mathbb{E}_{z_1,z_2}[f(z_1+z_2)g(z_1)h(z_2)]$ for some functions $g,h:\{0,1\}^n\to\{0,1\}$.

[Hint: Fix all the variables other than z_1 and z_2 such that that the expectation is maximized.]

(c) Use the result of Problem 1 to conclude that the expression in the above (for non-empty sums) is at most ε (i.e., $\mathbb{E}_{z_1,\ldots,z_k}\left[\prod_{(i,j)\in S} f(z_i)f(z_j)f(z_i+z_j)\right]\le\varepsilon$ for non-empty S).

(d) Conclude that $\Pr[\mathsf{acc}]$ is at most $2^{-\binom{k}{2}}+\varepsilon$.

3. **[Affine subspaces sample well]**

In the proof of the low-degree test, we assumed that affine subspaces are good samplers. In this problem, we will formally prove this statement.

Let $A\subset\mathbb{F}^m$ of density μ (i.e., $|A|=\mu q^m$).

$$\mathrm{Var}_{s\in\mathcal{S}_k^m}\left[\frac{|s\cap A|}{|s|}\right]\le\frac{\mu}{q}.$$

Hence, conclude that

$$\Pr_{s\in\mathcal{S}_k^m}\left[\left|\frac{|s\cap A|}{|s|}-\mu\right|\ge\varepsilon\right]\le\frac{\mu}{\varepsilon^2 q}.$$

4. **[polynomial decoding: short list of polynomials]**

Let $A:\mathbb{F}^m\to\mathbb{F}$ be any function (not necessarily a low degree polynomial). Let $p_1,p_2,\ldots,p_t:\mathbb{F}^m\to\mathbb{F}$ be the list of *all* degree d polynomials such that $\Pr_x[A(x)=p_i(x)]\ge\delta$. In other words, $p_1,\ldots,p_t$ is the list of *all* polynomials that have each agreement at least δ with the function A. Assume $\delta\ge2\sqrt{d/q}$. Prove that $t\le 2/\delta$. Hence, there are not too many low-degree polynomials that have considerable agreement with two polynomials.

[Hint: Use the fact that two low degree polynomial agree on at most d/q fraction of points (Schwartz-Zippel Lemma)]

5. **[Interpolation from cliques of consistency graph]**

   In lecture, we defined the notion of a consistency graph $G = (V, E)$, given a subspace oracle $A : \mathcal{S}_k^{k+1} \to P_{k,d}$ where $V = \mathcal{S}_k^m$ and $E = \{(s_1, s_2) | \forall x \in s_1 \cap s_2, A(s_1)(x) = A(s_2)(x)\}$. Suppose there exists a clique $W \subset V$ of size $\left(\frac{2d+1}{q}\right)|V|$, prove that there exists a polynomial $Q : F^m \to F$ of degree $2d$ such that for eah $w \in W$, we have $Q|_w \equiv A(w)$.

   [Hint: Use the large size of W to show that there exists two sets of d parallel hyperplanes (i.e, affine spaces of dimension k) in W. Interpolate along these hyperplanes to obtain a degree $2d$ polynomial Q. Use Schwartz-Zippel repeatedly to argue that Q identifies with $A(s)$ for all hyperplanes $s \in W$]

6. **[Degree reduction]**

   In lecture, we showed that if the plane-point low-degree test passes with with non-significant probability γ, in other words

$$\Pr_{s \in \mathcal{S}_k^m, x \in s} [A(s)(x) = A(x)] \geq \gamma,$$

   then there exists a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ of degree at most $2d$ such that

$$\Pr_x [Q(x) = A(x)] \geq \gamma^2 - \varepsilon,$$

   for some $\varepsilon = m^\alpha (d/q)^\beta$. In this problem, we will show that the degree of the polynomial Q can be reduced from $2d$ to d.

   Suppose there exists a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ of degree δq for some $0 < \delta < 1$ and furthermore,

$$\Pr_{s \in \mathcal{S}_k^m} [Q|_s \equiv A(s)] \geq \delta + \frac{1}{q},$$

   show that the degree of Q is in fact, at most d.

   [Hint: Suppose by contradiction this is not the case (i.e., degree$(Q) = D > d$. Consider any k dimensional affine subspace $s = z_0 + \text{span}\{z_1, z_2, \ldots, z_k\}$ for linearly independent $z_1, \ldots, z_k$. Any point in s is of the form $z_0 + \sum \alpha_i z_i$. Consider the coefficient of α_i^D in the polynomial $P(\alpha_1, \ldots, \alpha_k) = Q(z_0 + \sum \alpha_i z_i)$. Show using Schwartz-Zippel Lemma that with high probability this coefficient is not zero. Hence, with high probability $Q|_s$ is a degree D polynomial. Contradiction]

7. **[low degree testing to list of polynomials]**

   In lecture, we showed that if there is a list of low-degree polynomials that agrees with the space oracle then low-degree test theorem is true. In this problem, we will show the converse of this statement.

   Suppose there exists a function $f : (0, 1) \to (0, 1)$ such that the following is true.

   "[Low Degree Test Theorem] For every function $A : \mathbb{F}^m \to \mathbb{F}$ and $A : \mathcal{S}_k^m \to P_{m,d}$ that satisfies

$$\Pr_{s,x} [A(s)(x) = A(x)] \geq \gamma,$$

we have

$$\Pr_x [A(x) = Q(x)] \geq f(\gamma)$$

for some polynomial Q of degree at most d (end of Low Degree Test Theorem)"

(recall that we proved the above in lecture for the function $f(\gamma) = \gamma^2 - \varepsilon$)

Let $\varepsilon_0 = \sqrt{d/q}$ and $\delta \in (\varepsilon_0, 1)$. Set $\delta' = f(\delta - \varepsilon_0) - \varepsilon_0 \geq 2\varepsilon_0$. Prove that for any function $B : \mathbb{F}^m \to \mathbb{F}$, there exists a list of at most $t \leq 2/\delta'$ polynomials $Q_1, \ldots, Q_t : \mathbb{F}^m \to \mathbb{F}$ of degree at most d such that

$$\Pr_{s \in \mathcal{S}_k^m, x \in s} [B(s)(x) \neq B(x) \wedge (\exists i, Q_i|_s \equiv B(s))] \geq 1 - \delta.$$

You may assume the result of Problem 4. We will prove the above statement as follows. Suppose for contradiction that the statement if false.

Let $Q_1, Q_2, \ldots, Q_t$ be the list of polynomials that have at least δ' agreement with B. By Problem 4, $t \leq 2/\delta'$. Suppose the statement was false. Consider the following 3 events for a random $s \in \mathcal{S}_k^m$ and $x \in s$.

- $C : B(s)(x) = B(x)$
- $P : \exists i \in [t], B(x) = Q_i(x)$
- $Q : \exists i \in [t], B(s) \equiv Q_i|_s$

(a) Show that $\Pr[C \wedge \bar{S}] > \delta$. $\bar{S}$ denotes the event "not S"

(b) Argue using Schwartz-Zippel Lemma, $\Pr[C \wedge \bar{P} | \not{S}] \leq \varepsilon_0$.

(c) Conclude that $\Pr[C \wedge \bar{P}] > \delta - \varepsilon_0$.

(d) Construct a new oracle $B' : \mathbb{F}^m \to \mathbb{F}$ as follows: let Q' be an arbitrary polynomial of degree exactly $d+1$. Set $B'(x)$ to be $Q'(x)$ on all points x that satisfy P and $B(x)$ otherwise. Let the space oracle of B' be the same as that of B. Show that

$$\Pr\left[B'(s)(x) = B'(x)\right] > \delta - \varepsilon_0.$$

(e) Conclude from the low-degree test theorem that there exists a polynomial Q of degree at most d such that $\Pr[Q'(x) = Q(x)] \geq f(\delta - \varepsilon_0)$. Argue that Q and Q' are distinct polynomials and hence,

$$\Pr[B'(x) = Q(x) \wedge B'(x) \neq B(x)] \leq \Pr[Q'(x) = Q(x)] \leq \frac{d+1}{q} \leq \varepsilon_0.$$

(f) Argue that $\Pr[B(x) = Q(x) = B'(x)] \geq f(\delta - \varepsilon_0) - \varepsilon_0 = \delta'$.

(g) Conclude from above that there exists a $i \in [t]$ such that $Q \equiv Q_i$ (i.e., Q and Q_i are identical polynomials)

(h) Conclude that $\delta' \leq \Pr[B(x) = Q_i(x) = B'(x)] \leq \Pr[Q'(x) = Q(x)] \leq \varepsilon_0$, which is a contradiction.

8. **[Fourier interpretations]**

Let $f : \{0,1\}^n \to \mathbb{R}$ and write the Fourier expansion of f, $f = \sum_{S \subseteq [n]} \hat{f}(S)\chi_S$ where $\chi_S : \{0,1\}^n \to \{-1,1\}$ is defined as

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i},$$

and $\hat{f} : 2^{[n]} \to \mathbb{R}$ is defined as follows:

$$\hat{f}(S) = \langle f, \chi_S \rangle = \mathbb{E}\left[f(x)(-1)^{\sum_{i \in S} x_i}\right].$$

All probabilities and expectations in this question are with respect to the uniform product probability distribution on $\{0,1\}^n$.

(a) Given a set $S \subseteq [n]$, define $f^{\leq S} : \{0,1\}^n \to \mathbb{R}$ by

$$f^{\leq S} = \sum_{T : T \subseteq S} \hat{f}(T)\chi_T.$$

Note that $f^{\leq S}(x)$ actually only depends on the bits of x in S; call these bits x_S. Show that $f^{\leq S}(x_S)$ is equal to the expected value of f conditioned on the bits x_S (i.e., $f^{\leq S}(x_S) = \mathbb{E}_{y \in \{0,1\}^n}[f(y) | y_S = x_S]$ (The expectation is thus over the bits of x not in S.

(b) Suppose f's range is $\{-1,1\}$; i.e., f is a Boolean-valued function. We define the influence of the ith coordinate on f to be $\text{Inf}_i(f) = \Pr_x[f(x) \neq f(x^{(i)})]$, where $x^{(i)}$ denotes the string x with the ith bit flipped. This measures how sensitive f is to flipping the ith coordinate. Show that

$$\text{Inf}_i(f) = \sum_{S : i \in S} \hat{f}(S)^2.$$

(c) Again, suppose f is a Boolean-valued function. f is said to be monotone if $f(x) \leq f(y)$ whenever $x \geq y$. (By $x \geq y$ we mean $x_i \geq y_i$ for all i.) For example, the AND function which is given $\text{AND}(x,y) = 1 - 2xy$ is monotone. Similarly, OR, and Majority are also monotone functions; Parity is not monotone.

Show that if f is monotone then $\text{Inf}_i(f) = \hat{f}(\{i\})$ for each $i \in [n]$.

(d) Once more, suppose f is Boolean-valued. Suppose we pick $x \in \{0,1\}^n$ at random and then form a string $y \in \{0,1\}^n$ as follows: for each $i = 1 \ldots n$ independently, we set $y_i = x_i$ with probability ρ and set y_i to be a uniformly random bit with probability $1 - \rho$. The noise stability of f at ρ is defined to be

$$\text{Stab}_\rho(f) = 2\Pr[f(x) = f(y)] - 1,$$

a number in the range $[-1,1]$. This measures in some way how stable f is when you flip about $\frac{1}{2}(1-\rho)$ input bits. Show that

$$\text{Stab}_\rho(f) = \sum_{S \subseteq [n]} \hat{f}(S)^2 \rho^{|S|}.$$