CAS CS 591 B: Communication Complexity

Prof. Mark Bun | Fall 2019

Lecture Notes 19:

Pattern Matrix Method Continued

Reading.

- Sherstov, The Pattern Matrix Method

In this lecture, we'll prove a version of the Pattern Matrix Method lifting approximate degree to discrepancy.

Theorem 1. *Let $F = f \circ g_m^n$ where $g_m : \{-1,1\}^m \times ([m] \times \{-1,1\})$ is given by $g_m(x,(i,w)) = x_i w$. Then for every $\delta > 0$,*

$$\mathrm{disc}(F) \leq \delta + m^{-\operatorname{adeg}_{1-\delta}(f)/2}.$$

Here, we recall the definition of approximate degree and its dual characterization.

Definition 2. Let $\varepsilon > 0$ and let $f : \{-1,1\}^n \to \{-1,1\}$. A real polynomial $p : \{-1,1\}^n \to \mathbb{R}$ ε-approximates f if $|p(x) - f(x)| \leq \varepsilon$ for all $x \in \{-1,1\}^n$. The ε-approximate degree of f is the least degree of a polynomial p which approximates f, and is denoted $\operatorname{adeg}_\varepsilon(f)$.

Theorem 3. *Let $f : \{-1,1\}^n \to \{-1,1\}$ be a boolean function. Then $\operatorname{adeg}_\varepsilon(f) > d$ if and only if there exists a function $\psi : \{-1,1\}^n \to \mathbb{R}$ such that*

1. $\langle f, \psi \rangle := \sum_{x \in \{-1,1\}^n} f(x)\psi(x) > \varepsilon$
2. $\|\psi\|_1 := \sum_{x \in \{-1,1\}^n} |\psi(x)| = 1$
3. $\hat{\psi}(S) = 2^{-n} \sum_{x \in \{-1,1\}^n} \psi(x)\chi_S(x) = 0$ *for every* $S \subseteq [n]$ *with* $|S| \leq d$. *Here* $\chi_S(x) = \prod_{i \in S} x_i$.

1 Pattern Matrix Method Proof

It will be convenient to define the notion of a "Pattern Matrix" of a function $\psi : \{-1,1\}^n \to \mathbb{R}$. In the special case where ψ is boolean, this is simply the communication matrix of $\psi \circ g_m^n$, but the definition of course also makes sense when ψ is real-valued.

Definition 4. The m-Pattern Matrix of a function $\psi : \{-1,1\}^n \to \mathbb{R}$ is the $2^{nm} \times (2m)^n$ real matrix $\mathrm{PM}_m(\psi)$ given by

$$\begin{aligned}\mathrm{PM}_m[(x_1,\ldots,x_n),((i_1,w_1),\ldots(i_n,w_n))] &= \psi(g(x_1,(i_1,w_1)),\ldots,g(x_n,(i_n,w_n))) \\ &= \psi(x|_I \oplus w)\end{aligned}$$

where $I = (i_1,\ldots,i_n)$ and the notation $x|_I$ indicates the projection of x to the coordinates specified by I, i.e., $(x_{1,i_1},\ldots,x_{n,i_n})$.

Every formulation of the Pattern Matrix Method makes use of the following lemma which relates the spectral norm of a pattern matrix to the Fourier coefficients of the underlying function.

Lemma 5. *Let $\psi : \{-1,1\}^n \to \mathbb{R}$ and let $\Psi = \mathrm{PM}_m(\psi)$ be its pattern matrix. Then the spectral norm of Ψ is given by*

$$\|\Psi\| = \sqrt{2^{nm} \cdot (2m)^n} \cdot \max_{S \subseteq [n]} \left(|\hat{\psi}(S)| \cdot m^{-|S|/2} \right).$$

Let us see how to use Lemma 5 to prove Theorem 1.

Proof of Theorem 1. Recall from our discussion of discrepancy that for any function $F : X \times Y \to \{-1,1\}$ we have

$$\mathrm{disc}(F) \le \frac{\|F\|}{\sqrt{|X||Y|}}.$$

(Here, for convenience, we will conflate a two-party function with its sign matrix.) The proof of this actually gives an upper bound on the discrepancy of F with respect to the uniform distribution. It can be generalized as follows. Let P be any matrix with non-negative entries which sum to 1. Then

$$\mathrm{disc}(F) \le \|F \circ P\| \cdot \sqrt{|X||Y|}$$

where $F \circ P$ is the matrix obtained by taking the entrywise product of F and P. This upper bound on discrepancy follows from the calculation

$$\begin{aligned}
\mathrm{disc}_P(F) &= \max_{S \subseteq X, T \subseteq Y} \left| \sum_{x \in S} \sum_{y \in T} P[x,y] F[x,y] \right| \\
&= \max_{S,T} |1_S^T (P \circ F) 1_T| \\
&\le \max_{S,T} \|1_S\|_2 \cdot \|P \circ F\| \cdot \|1_T\|_2 \\
&= \|P \circ F\| \sqrt{|X||Y|}.
\end{aligned}$$

Now suppose $f : \{-1,1\}^n \to \{-1,1\}$ is such that $\deg_{1-\delta}(f) > d$. Let Ψ be the $2^{nm} \times (2m)^n$ Pattern Matrix of $2^{-nm} \cdot m^{-n} \cdot \psi$. Then we have $\|\Psi\|_1 = 1$ and $\langle \Psi, S_F \rangle > 1 - \delta$ by Theorem 3. Now we calculate $\|\Psi\|$. First observe that for every $S \subseteq [n]$,

$$|\hat{\psi}(S)| = 2^{-n} \left| \sum_x \psi(x) \chi_S(x) \right| \le 2^{-n} \|\psi\|_1 = 2^{-n}.$$

Using the fact that $\hat{\psi}(S) = 0$ for all $|S| \le d$, we have by Lemma 5 that

$$\|\Psi\| \le \sqrt{s} \cdot (2^{-nm} \cdot m^{-n}) \cdot 2^{-n} m^{-d/2} = s^{-1/2} m^{-d/2}.$$

where $s = 2^{nm} \cdot (2m)^n$ is the size of Ψ.

Now let us write $\Psi = P \circ H$ where P is a non-negative matrix whose entries sum to 1 and H is a sign matrix. We can do this because $\|\Psi\|_1 = 1$. The above discrepancy calculation then shows that

$$\mathrm{disc}_P(H) \le \|P \circ H\| \sqrt{s} \le m^{-d/2}.$$

Moreover, applying the triangle inequality to the definition of discrepancy,

$$\mathrm{disc}_P(F) \leq \mathrm{disc}_P(H) + \|(F-H)\circ P\|_1.$$

Let $E = \{(x,y) : F(x) \neq H(x)\}$. We can equivalently write the error term as

$$\begin{aligned}
\|(F-H)\circ P\|_1 &= 2\sum_E P(x,y) \\
&= \sum_{(x,y)\in \bar{E}} P(x,y) + \sum_{(x,y)\in E} P(x,y) - \left(\sum_{(x,y)\in \bar{E}} P(x,y) - \sum_{(x,y)\in E} P(x,y)\right) \\
&= 1 - \langle F, H\circ P\rangle \\
&= 1 - \langle F, \Psi\rangle \\
&\leq 1 - (1-\delta).
\end{aligned}$$

Putting everything together, we conclude

$$\mathrm{disc}_P(F) \leq \mathrm{disc}_P(H) + \|(F-H)\circ P\|_1 \leq m^{-d/2} + \delta.$$

□

2 Proof of Lemma 5

We now prove Lemma 5 relating the spectral norm of a pattern matrix $\mathrm{PM}_m(\psi)$ to the Fourier coefficients of ψ.

A key fact about the Fourier representation of ψ is that we have

$$\psi(x) = \sum_{S\subseteq[n]} \hat{\psi}(S)\chi_S(x).$$

For each $S \subseteq [n]$ let A_S be the pattern matrix $\mathrm{PM}_m(\chi_S)$. Then by linearity,

$$\Psi = \mathrm{PM}_m(\psi) = \sum_{S\subseteq[n]} \hat{\psi}(S) A_S.$$

To understand the singular values of Ψ, we invoke the following lemma relating the singular values of a sum of matrices to the singular values of the individual matrices.

Lemma 6. *Let A and B be real matrices with $AB^T = 0$ and $A^TB = 0$. Then the multiset of nonzero singular values of $A+B$ is the union of the singular values of A with singular values of B.*

We won't prove the lemma, but the idea is as follows. The singular values of $A+B$ are just the square roots of the eigenvalues of $(A+B)(A+B)^T = AA^T + BB^T$. The orthogonality of A and B further implies that vectors in the spectral decomposition of AA^T are orthogonal to those in the spectral decomposition of BB^T. Hence the set of eigenvalues of $AA^T + BB^T$ is just the union of the eigenvalues of AA^T and BB^T.

In order to apply the lemma, we need to show that the matrices A_S are orthogonal. To see this, let $S, T \subseteq [n]$ with $S \neq T$. Then for every $x, x' \in \{-1,1\}^{nm}$,

$$
\begin{aligned}
A_S A_T^T[x,x'] &= \sum_I \sum_w \chi_S(x|_I \oplus w)\chi_T(x'|_I \oplus w) \\
&= \sum_I \chi_S(x|_I)\chi_T(x|_I) \sum_w \chi_S(w)\chi_T(w) \\
&= 0
\end{aligned}
$$

because χ_S and χ_T are orthogonal. A similar argument can be used to show that

$$A_S^T A_T = 0.$$

So by the lemma, the set of nonzero singular values of Ψ is just the union of the nonzero singular values of the matrices $\hat{\psi}(S)A_S$. We will be done if we can show that the only nonzero eigenvalue of $A_S^T A_S$ is $2^{nm} \cdot (2m)^n \cdot m^{-|S|}$ (with multiplicity $m^{|S|}$).

This can be done by writing $A_S^T A_S = W \otimes V$ where $W \in \{-1,1\}^{2^n \times 2^n}$ is given by

$$W[w,w'] = \chi_S(w)\chi_S(w')$$

and $V \in \mathbb{R}^{m^n \times m^n}$ is

$$V[I,I'] = \sum_{x \in \{-1,1\}^n} \chi_S(x|_I)\chi_S(x_{I'}).$$

The first matrix W has rank 1 and it is easy to see that it has 2^n as its only singular value. The second matrix V is similar to $2^{mn} \operatorname{diag}(J,\ldots,J)$ where J is the all-ones square matrix with $m^{n-|S|}$ rows. Hence the only nonzero singular value of V is $2^{mn} \cdot m^{n-|S|}$ (with multiplicity $m^{|S|}$). So the only nonzero eigenvalue of $A_S^T A_S$ is $2^{nm} \cdot (2m)^n \cdot m^{-|S|}$.

Now Lemma 5 follows because the spectral norm of Ψ is the largest singular value of any of the matrices $\hat{\psi}(S)A_S$.