Dictatorship testing for Max Cut

Venkatesan Guruswami, June 2011, Fields Institute

We now turn to the simplest binary CSP, Max Cut.

As we saw, there is a 0.878.. approx. alg. for Max Cut. Is this optimal?
(For $c \geq 0.845$, the precise guarantee was $\frac{\cos^{-1}(1-2c)}{\pi}$)

Given the earlier discussion, let us focus on giving a dictatorship test for Max Cut.

Recall goal: Given $f: \{0,1\}^R \to \{1,-1\}$ (let's focus on ±1 valued functions for simplicity)

test if f = dictator ("projection")

or far from dictator (has no "influential coordinate")

with a cut test, i.e. pick $(x,y) \in D$ (suitable distribution D)

and check $f(x) \neq f(y)$

How should we pick x, y?

$\forall i, \quad \Pr[x_i \neq y_i] \geq c$ (we will have $c > 1/2$)

So let's pick $x \in \{0,1\}^R$ u.a.r

and $\forall i$, $y_i = x_i$ with prob $1-c$

$= \overline{x_i}$ with prob c

$y = N_c(x)$ (noisy version of x)

Clearly, $\Pr_{(x,y)}[f(x) \neq f(y)] = c$

if $f(x) = x_i$.

Soundness: What about f s.t. $\mathrm{Inf}_i(f)$ is small ($< \tau$) $\forall i$?

$\Pr[f(x) \neq f(y)] = \text{Noise-Sensitivity}_c(f)$

Exer: $\text{Noise-sensitivity}_c(f) = \dfrac{1 - \sum \hat{f}(\alpha)^2 (1-2c)^{|\alpha|}}{2}$.

To get some intuition, let us consider the quintessential function without influential coordinates: Majority

What is $\Pr[\mathrm{Maj}(x) \neq \mathrm{Maj}(y)]$?

It is convenient to change domain to $\{1,-1\}^R$, so $\mathrm{Maj}(x) = \mathrm{sgn}\left(\sum x_i\right)$

We are interested in $\Pr\left[\operatorname{sgn}\left(\frac{1}{\sqrt{R}}\sum x_i\right) \neq \operatorname{sgn}\left(\frac{1}{\sqrt{R}}\sum y_i\right)\right]$

$\frac{1}{\sqrt{R}}\sum x_i \longrightarrow N(0,1)$ by the Central Limit Thm

Also $\frac{1}{\sqrt{R}}\sum x_i \to \frac{1}{\sqrt{R}}\sum g_i$ [g_i = i.i.d $N(0,1)$ r.v's.]

Likewise $\frac{1}{\sqrt{R}}\sum y_i \to \frac{1}{\sqrt{R}}\sum \mu_i g_i$

($\mu_i = 1$ with prob $1-c$, -1 with prob. c)

Let $a = \left(\frac{1}{\sqrt{R}}, \ldots \frac{1}{\sqrt{R}}\right)^T$, $b = \left(\frac{\mu_1}{\sqrt{R}}, \ldots \frac{\mu_R}{\sqrt{R}}\right)^T$

a, b are unit vectors.

Desired prob $= \Pr_{g \sim N(0,1)^R}\left[\operatorname{sgn}(a \cdot g) \neq \operatorname{sgn}(b \cdot g)\right]$

(Recall GW analysis?)

$= \frac{\angle a,b}{\pi} = \frac{\cos^{-1}(a \cdot b)}{\pi} = \frac{\cos^{-1}(1-2c)}{\pi}$

Thus f = Majority, passes test with prob $\approx \frac{\cos^{-1}(1-2c)}{\pi}$.

Is this the worst case?

Theorem (Majority is Stablest) (MOO 2005)

(actually 'most sensitive' (for $c > 1/2$))

For $c > 1/2$, following is true.

$\forall \epsilon > 0 \ \exists \tau > 0$ s.t. if $\mathrm{Inf}_i(f) < \tau \ \forall i$,

then $\text{Noise-Sensitivity}_c(f) \leq \frac{\cos^{-1}(1-2c)}{\pi} + \epsilon$

This gives a c vs $\frac{\cos^{-1}(1-2c)}{\pi} + \epsilon$ dictatorship test for Max Cut

Cor: The 0.878.. factor of [GW] is best possible, assuming the UGC!

Proof of 'maj. is stablest' proceeds via "invariance principle" which is a higher-degree generalization of the central limit theorem.

Thm (somewhat informal): Let $X = (X_1 \ldots X_n)$ be uniform from $\{1, -1\}^n$ & let $Q(X_1 \ldots X_n)$ be a degree d multilinear polynomial ($\mathrm{Var}[Q] = 1$ & $\mathrm{Inf}_i(Q) \leq \tau \ \forall i$)

Then, $\sup_t \left[\left| \Pr[Q(X_1 \ldots X_n) \leq t] - \Pr[Q(G_1 \ldots G_n) \leq t] \right| \right]$

$\leq O(d \tau^{1/8d})$

This form is due to (MOO, 2005). Related 'qualitative' form in (Rötar, 79)

[Raghavendra '08] Generic conversion of an integrality gap for [GW] SDP for Max Cut to matching dictatorship test.

Dict test

$\mathrm{SDPval}(G) = c$

$\mathrm{MaxCut}(G) = s$

$\Longrightarrow$ $\begin{cases} \Pr[F(x) \neq F(y)] \geq c - \varepsilon & \text{if } F = \text{dictator} \\ \Pr[F(x) \neq F(y)] \leq s + \varepsilon & \text{if } F \text{ far from dictator} \end{cases}$

More generally, he showed such a result for every CSP:

c vs s gap for canonical/simple SDP

$\Longrightarrow$ c vs s dictatorship test

The [KKMO] test obtained when the [Feige-Schechtman] integrality gap for SDP used in this framework.

A crucial technical tool used in Raghavendra's analysis is a more general invariance principle by Mossel.

Let's give a glimpse of this in a special case.

Suppose $f: \{1,-1\}^R \to \{1,-1\}$ is such that

$$\mathrm{Inf}_i(f) \leq \tau \text{ (small) } \forall i$$

Let D be a distribution on $\{1,-1\}^k$ is pairwise independent (i.e. the marginal distribution on Y_{i_1} & Y_{i_2} for $(Y_1 \ldots Y_k) \sim D$ is uniform, $\forall i_1, i_2$)

Let $X = \begin{pmatrix} \text{———} X_1 \text{———} \\ \text{———} X_2 \text{———} \\ \vdots \\ \text{———} X_k \text{———} \end{pmatrix}$ ↳ columns drawn from D^R

Then

$$\mathop{E}_X\left[\prod_{i=1}^{k} f(X_i)\right] \underset{\epsilon}{\approx} \prod_{i=1}^{k} E[f(X_i)]$$

(i.e. can replace X_i's by independent random strings !!)

In words, low-influence functions cannot distinguish distributions with matching first & second moments

More generally, the expectation $\mathbb{E}\left[\prod_{i=1}^{k} f(x_i)\right]$ is preserved (up to small errors) if distribution D is replaced by any distribution Δ on $\mathbb{R}^k$ s.t.

$$\mathop{\mathbb{E}}_{Y\sim D}[Y_i] = \mathop{\mathbb{E}}_{Z\sim\Delta}[Z_i] \; ; \; \mathop{\mathbb{E}}_{Y\sim D}[Y_i Y_j] = \mathop{\mathbb{E}}_{Z\sim\Delta}[Z_i Z_j]$$

when f has no influential coordinate.

This feature is used (at a very high level) as follows. The vectors in the SDP solution are used to sample from a global distribution (playing the role of Δ) with matching 1st & 2nd moments with D, and this is used to "round" the vectors to a solution with value $\approx$ the prob. that f passes the dictatorship test. ☒