Part C Probabilistic Combinatorics 2008

Remarks on Azuma, concentration of Lipschitz functions, and isoperimetric inequalities

When doing the isoperimetric inequality in lectures, I cheated inadvertently by using a stronger version of the result on concentration of Lipschitz functions than I had previously proved.

Here are a few notes to clarify things for those who are interested.

We proved the following Lemma:

Lemma 1 *Let $\alpha > 0$. If $|Y| \leq 1$ with probability 1 and $\mathbb{E}\, Y = 0$, then $\mathbb{E}\left(e^{\alpha Y}\right) \leq e^{\alpha^2/2}$.*

From this we deduced the following form of Azuma's inequality:

Theorem 1 *Let $0 = X_0, X_1, \dots, X_m$ be a martingale such that $|X_{i+1} - X_i| \leq 1$ with probability 1, for $0 \leq i \leq m-1$. Then for any $t > 0$,*

$$\mathbb{P}\left(X_m > t\right) \leq \exp\left(-\frac{t^2}{2m}\right).$$

This in turn gave us the concentration result for Lipschitz functions:

Theorem 2 *Let $V_1, \dots, V_m$ be independent random variables taking values in sets $\Omega_1, \dots, \Omega_m$ respectively. Let $f : \prod_{i=1}^m \Omega_i \to \mathbb{R}$ be a Lipschitz function. Let $X = f(V_1, \dots, V_m)$. Then*

$$\mathbb{P}\big(|X - \mathbb{E}\, X| \geq t\big) \leq 2\exp\left(-\frac{t^2}{2m}\right).$$

The idea of the proof was to consider the martingale (X_i) where $X_i = \mathbb{E}\left(X | V_1, \dots, V_i\right)$. Since the ith coordinate cannot affect the function f by more than 1, learning the value of V_i does not affect the conditional expectation of X by more than 1. Hence $|X_{i+1} - X_i| \leq 1$ with probability 1, and we can apply Azuma's inequality.

In fact, more is true. Once we know $V_1, \dots, V_i$, the possible *range* of values of X_{i+1} is at most 1; that is, given $V_1, \dots, V_i$, there exist a and b such that $b - a \leq 1$ and such that $a \leq X_{i+1} \leq b$ with probability 1.

This observation makes it possible to prove a stronger version of Theorem 2. We use the following generalisation of Lemma 1:

Lemma 2 *Let $\alpha > 0$. If $a \leq Y \leq b$ with probability 1, and $\mathbb{E}\, Y = 0$, then $\mathbb{E}\left(e^{\alpha Y}\right) \leq e^{\alpha^2 (b-a)^2/8}$.*

Using this one can prove a more general form of Azuma's inequality which leads to the following stronger version of Theorem 2:

Theorem 3 *Let $V_1, \ldots, V_m$ be independent random variables taking values in sets $\Omega_1, \ldots, \Omega_m$ respectively. Let $f : \prod_{i=1}^m \Omega_i \to \mathbb{R}$ be a Lipschitz function. Let $X = f(V_1, \ldots, V_m)$. Then*

$$\mathbb{P}\big(|X - \mathbb{E}\, X| \geq t\big) \leq 2 \exp\left(-\frac{2t^2}{m}\right).$$

This is the result I was claiming during the proof of the isoperimetric inequality. Namely, we used Theorem 3 to show:

Theorem 4 *Let $X = (X_1, X_2, \ldots, X_n)$ be a vector with independent entries $X_i \in \Omega_i$. Let $A \subset \Omega = \prod_{i=1}^n \Omega_i$. Then for any $t \geq 0$,*

$$\mathbb{P}\big(X \in A\big)\mathbb{P}\big(d_H(X, A) \geq t\big) \leq \exp\left(-\frac{t^2}{2n}\right),$$

(where d_H is the Hamming distance).

If one instead uses the weaker version in Theorem 2, one can get a similar result but with $\exp\left(-\frac{t^2}{8n}\right)$ on the right instead. (Exercise: prove this using the same method from lectures).

To deduce Theorem 3 from 2, one needs a more general version of Azuma's inequality which needs a slightly more general definition of a martingale than the one we used in the course. The argument is quite straightforward though.

For full details of all of this (and much more beyond), a good place to look is Colin McDiarmid's article "Concentration" in the book *Probabilistic Methods for Algorithmic Discrete Mathematics*, ed. M. Habib et al, Springer 1998.

Course webpage: `http://www.stats.ox.ac.uk/~martin/PC.html`