Rounding Techniques in Approximation Algorithms

Lecture 17: Martingales

Lecturer: Nathan Klein *Scribe: Tolik Zinovyev*

1 Introduction to Martingales

We introduce the concept of a martingale and prove a crucial lemma that will lead to showing that sub-isotropic updates give us concentration.

Martingales are sequences of random variables that find applications in random walk and gambling problems. A rich collection of Chernoff-like tail bounds are known for martingales which make them a useful tool. Martingales are defined as follows:

Definition 1.1 (Martingale)**.** *A sequence of random variables $Z_0, Z_1, ...$ is a martingale with respect to a sequence of random variables $X_0, X_1, ...$ if and only if for all $n \geq 0$ the following are true.*

- *Z_n is a function of $X_0, X_1, ..., X_n$;*
- *$\mathbb{E}[|Z_n|] < \infty$, i.e., Z_n has well defined expectation;*
- *$\mathbb{E}[Z_{n+1}|X_0, ..., X_n] = Z_n$.*

The last condition is something we saw in pipage rounding: the expectation of i-th coordinate after rounding conditioned on its present value is equal to its present value. One can think of variables $X_0, X_1, ...$ as new information about the state of the world / experiment, and Z_i as some running metric we care about.

The following is an example of a martingale. Suppose a gambler plays a sequence of fair games. Define X_i to be his gain in game i (which can be negative), and Z_i to be the gambler's total gains after i games. More formally, we have $X_0, X_1, ...$ with $\mathbb{E}[X_{n+1}|X_0, ..., X_n] = 0$ and $Z_0, Z_1, ...$ with $Z_n = \sum_{i=0}^{n} X_i$. To prove $Z_0, Z_1, ...$ is a martingale with respect to $X_0, X_1, ...$, let us show the third condition, as the first two are obvious:

$$\begin{aligned}
\mathbb{E}[Z_{n+1}|X_0, ..., X_n] &= \mathbb{E}[Z_n + X_{n+1}|X_0, ..., X_n] \\
&= \mathbb{E}[Z_n|X_0, ..., X_n] + \mathbb{E}[X_{n+1}|X_0, ..., X_n] \\
&= \mathbb{E}[Z_n|X_0, ..., X_n] \\
&= Z_n
\end{aligned}$$

where the last equality holds since Z_n is a function of $X_0, ..., X_n$. It is important to note that the bets can have different amounts and can even depend on the outcomes of the previous games and $Z_0, \ldots,$ would still be a martingale.

A special type of martingale is a Doob martingales and is constructed as follows. Let $X_0, X_1, ..., X_n$ be random variables, and let Z be a random variable with $\mathbb{E}[|Y|] < \infty$ (Y will generally depend on $X_0, ..., X_n$). Also define $Z_i = \mathbb{E}[Y|X_0, ..., X_i]$ for all $0 \leq i \leq n$. Then $Z_0, ..., Z_n$ is a martingale with respect to $X_0, ..., X_n$ since

$$\begin{aligned}
\mathbb{E}[Z_{i+1}|X_0, ..., X_i] &= \mathbb{E}\big[\mathbb{E}[Y|X_0, ..., X_{i+1}]\big|X_0, ..., X_i\big] \\
&= \mathbb{E}[Y|X_0, ..., X_i] \\
&= Z_i.
\end{aligned}$$

Consider the following example of a Doob martingale. We have m balls that we throw uniformly at random in one of n bins. We care about the number of empty bins after all m balls are thrown. Let Y be this random number, and let X_i for $1 \leq i \leq m$ be the bin where ball i lands (we can define $X_0 = -1$). Then Z_0 is simply $\mathbb{E}[Y|X_0] = \mathbb{E}[Y]$, a deterministic number. However, for $i \geq 1$, each $Z_i = \mathbb{E}[Y|X_0, ..., X_i]$ is a refined estimate of the eventual number of empty bins as a function of the "known" outcomes $X_1, ..., X_i$. As i increases, Z_i becomes a more precise estimate of Y.

Many Chernoff like tail bounds are known for martingales, here is a well-known one.

Theorem 1.2 (Azuma-Hoeffding Inequality)**.** *Let $X_0, ..., X_n$ be a martingale (with respect to itself) such that $X_k - X_{k-1} \leq c_k$. Then, for all $t \geq 1$ and $\lambda > 0$,*

$$\Pr\left[|X_t - X_0| \geq \lambda\right] \leq 2 \cdot \exp\left(-\frac{\lambda^2}{2\sum_{k=1}^{t} c_k^2}\right).$$

For instance, one can use it to upper bound $\Pr\left[|Y - \mathbb{E}[Y]| \geq \epsilon\right]$ for the balls and bins problem. $\mathbb{E}[Y]$ doesn't even need to be known.

2 Application to Sub-isotropic Rounding

Ideally, we would apply Theorem 1.2 to show concentration, but it's somewhat difficult to apply to sub-isotropic rounding. Even though we have $Y_k = X_k - X_{k-1} \leq 1$, our random walk makes very small updates (recall our updates were $\epsilon U^{1/2} r$ for $\epsilon \leq n^{-3/2}/2$), so the N for which X_N is integral may be fairly large compared to $\mathbb{E}[X]$. This is why we do the more involved Freedman-type bound as below which does not depend on N. We will use this lemma to complete the proof that sub-isotropic updates imply concentration.

Lemma 2.1. *Let $0 < \alpha < 1$, $t \geq 0$. Let $Z_0, Z_1, ...$ be random variables with Z_0 deterministic. Let $Y_k = Z_k - Z_{k-1} \leq 1$. Finally, assume*

$$\mathbb{E}[Y_k|Z_1, ..., Z_{k-1}] \leq -\alpha \mathbb{E}[Y_k^2|Z_1, ..., Z_{k-1}].$$

Then

$$\Pr[Z_k - Z_0 > t] \leq e^{-\alpha t}.$$

Notice that, unlike in Azuma-Hoeffding Inequality, the right hand side of the tail bound inequality doesn't feature k. The price of it is that we don't assume a martingale: instead of $\mathbb{E}[Y_k|Z_1, ..., Z_{k-1}] = 0$, we have $\mathbb{E}[Y_k|Z_1, ..., Z_{k-1}] \leq -\alpha \mathbb{E}[Y_k^2|Z_1, ..., Z_{k-1}] < 0$, i.e. we must skew in the negative direction as time evolves.

Proof. By Markov's inequality, we have

$$\Pr[Z_k - Z_0 > t] = \Pr[e^{\alpha(Z_k - Z_0)} \geq e^{\alpha t}] \leq \frac{\mathbb{E}[e^{\alpha(Z_k - Z_0)}]}{e^{\alpha t}}.$$

This is at most $e^{-\alpha t}$ if and only if $\mathbb{E}[e^{\alpha(Z_k - Z_0)}] \leq 1$ if and only if $\mathbb{E}[e^{\alpha Z_k}] \leq e^{\alpha Z_0}$ since Z_0 is deterministic.

Denote $\mathbb{E}_{k-1}[\,\cdot\,] = \mathbb{E}[\,\cdot \mid Z_1, ..., Z_{k-1}]$. We now bound the following.

$$\mathbb{E}_{k-1}[e^{\alpha Z_k}] = e^{\alpha Z_{k-1}} \cdot \mathbb{E}_{k-1}\left[e^{\alpha(Z_k - Z_{k-1})}\right] = e^{\alpha Z_{k-1}} \cdot \mathbb{E}_{k-1}\left[e^{\alpha Y_k}\right] \; [\leq]$$

by the below computation (Lemma 2.2), this is at most

$$[\leq] \; e^{\alpha Z_{k-1}} \cdot \exp\left(\alpha \mathbb{E}_{k-1}[Y_k] + (e^{\alpha} - \alpha - 1)\mathbb{E}_{k-1}[Y_k^2]\right) \leq$$
$$e^{\alpha Z_{k-1}} \cdot \exp\left((e^{\alpha} - \alpha^2 - \alpha - 1)\mathbb{E}_{k-1}[Y_k^2]\right)$$

This is at most $e^{\alpha Z_{k-1}}$ since $e^{\alpha} \leq 1 + \alpha + \alpha^2$ for $0 \leq \alpha \leq 1$.

So, we want to show $\mathbb{E}[e^{\alpha Z_k}] \leq e^{\alpha Z_0}$ and we have $\mathbb{E}_{k-1}[e^{\alpha Z_k}] \leq e^{\alpha Z_{k-1}}$. Applying $\mathbb{E}_{k-2}[\,\cdot\,]$ to both sides, we have

$$\mathbb{E}_{k-2}\left[e^{\alpha Z_k}\right] = \mathbb{E}_{k-2}\left[\mathbb{E}_{k-1}\left[e^{\alpha Z_k}\right]\right] \leq \mathbb{E}_{k-2}\left[e^{\alpha Z_{k-1}}\right] \leq e^{\alpha Z_{k-2}}.$$

Repeating this $k-2$ more times, we have

$$\mathbb{E}\left[e^{\alpha Z_k}\right] = \mathbb{E}_0\left[e^{\alpha Z_k}\right] \leq e^{\alpha Z_0}.$$

as desired. □

Finally, we show the computation used above.

Lemma 2.2. *If $X \leq 1$ and $\lambda > 0$, then $\mathbb{E}[e^{\lambda X}] \leq \exp\left(\lambda \mathbb{E}[X] + (e^{\lambda} - \lambda - 1)\mathbb{E}[X^2]\right)$.*

Proof. Define

$$g(x) = \begin{cases} (e^x - x - 1)/x^2 & \text{if } x \neq 0 \\ 1/2 & \text{if } x = 0. \end{cases}$$

It can be seen that $g(x)$ is increasing for all $x \in \mathbb{R}$. Then $e^x - x - 1 = g(x)x^2$ and $e^x = 1 + x + g(x)x^2$.

When $x \leq 1$,

$$e^{\lambda x} =$$
$$1 + \lambda x + g(\lambda x) \cdot (\lambda x)^2 \leq$$
$$1 + \lambda x + g(\lambda) \cdot (\lambda x)^2 =$$
$$1 + \lambda x + (e^{\lambda} - \lambda - 1)x^2.$$

Thus,

$$\mathbb{E}[e^{\lambda X}] \leq$$
$$1 + \lambda \mathbb{E}[X] + (e^{\lambda} - \lambda - 1)\mathbb{E}[X^2] \leq$$
$$\exp\left(\lambda \mathbb{E}[X] + (e^{\lambda} - \lambda - 1)\mathbb{E}[X^2]\right).$$

□