Mathematical Toolkit	**Autumn 2015**

Lecture 14: November 16, 2015

Lecturer: Madhur Tulsiani

1 The power of two random choices

We will now show that two random choices can reduce the maximum load to $O(\ln \ln n)$. The proof technique is due to Azar et al. [ABKU94, ABKU99] and various applications were explored by Mitzenmacher in his thesis [Mit96]. We first provide the intuition for the proof.

For each i, let B_i denote the number of bins with at least i balls. Suppose $B_i \leq \beta_i$ for some bound β_i. Then B_{i+1} is bounded above by a binomial random variable corresponding to the number of heads in n independent coin tosses, where the probability of each toss being heads is at most $(\beta_i/n)^2$. This is because for a ball to land a bin such that the load of the bin becomes greater than i, it must happen that both the random bins which we chose to put it in, had load at least i. This happens with probability at most $(\beta_i/n)^2$. Thus, B_{i+1} is upper bounded by the above random variable, which we denote as $\mathsf{Bin}\left(n, \left(\frac{\beta_i}{n}\right)^2\right)$.

This, $\mathbb{E}\left[B_{i+1}\right] \leq n \cdot \left(\frac{\beta_i}{n}\right)^2$ and B_{i+1} is at most $e \cdot \frac{\beta_i^2}{n}$ with high probability. We can then take β_{i+1} to be $e \cdot \frac{\beta_i^2}{n}$. For the above sequence, the value of β_i becomes less than 1 for $i_0 = O(\ln \ln n)$, and thus we can bound the maximum load by i_0. The proof will follow this intuition, except that for the last step, when $\mathbb{E}\left[B_i\right]$ becomes very small, we will not be able to use a Chernoff bound and will have to resort to a slightly different analysis.

We first define the values β_i. Let $\beta_6 = \frac{n}{2e}$ and $\beta_{i+1} = e \cdot n \cdot \left(\frac{\beta_i}{n}\right)^2$.

$$\begin{aligned}
& \beta_6 = \frac{n}{2e} \\
\Rightarrow \quad & \beta_7 = e\left(\frac{n}{2e}\right)^2 n = \frac{n}{4e} = \frac{n}{2^2 e} \\
\Rightarrow \quad & \beta_8 = e\left(\frac{n}{4e}\right)^2 n = \frac{n}{16e} = \frac{n}{2^{2^2} e} \\
\Rightarrow \quad & \beta_9 = e\left(\frac{n}{16e}\right)^2 n = \frac{n}{256e} = \frac{n}{2^{2^3} e} \\
& \vdots \\
\Rightarrow \quad & \beta_i = \frac{n}{2^{2^{i-6}} e}
\end{aligned}$$

Let E_i be the event that $B_i \leq \beta_i$. Note that E_6 holds for sure since there can be at most $n/6 \leq n/2e$ bins with 6 or more balls. We show that with high probability, if E_i holds then E_{i+1} holds provided $\beta_i^2 \geq 2n \ln n$.

Claim 1.1 *Let i be such that $\beta_i^2 \geq 2n \ln n$. Then,*

$$\mathbb{P}\left[\neg E_{i+1} \mid E_i\right] \;\leq\; \frac{1}{n^2} \cdot \frac{1}{\mathbb{P}\left[E_i\right]}.$$

Proof: The tricky part in proving the claim is the conditioning. Conditioning on the event E_i, the choices made by the various balls are no longer independent. To take care of this, we define the random variables Y_t as

$$Y_t \;=\; \begin{cases} 1 & \text{if at time } t \text{ there are at most } \beta_i \text{ bins with load } i \text{ and} \\ & \text{both bins chosen by the } t^{th} \text{ ball have load at least } i \\ 0 & \text{otherwise} \end{cases}$$

We can now write the event E_{i+1} in terms of the variables Y_t. We have

$$\mathbb{P}\left[\neg E_{i+1} \mid E_i\right] \;=\; \frac{\mathbb{P}\left[\sim E_{i+1} \wedge E_i\right]}{\mathbb{P}\left[E_i\right]} \;\leq\; \frac{\mathbb{P}\left[\sum_{t=1}^{n} Y_t \geq \beta_{i+1}\right]}{\mathbb{P}\left[E_i\right]}$$

Note that the variables Y_t are still *not* independent, but satisfy that

$$\mathbb{P}\left[Y_t = 1 \mid Y_1, \ldots, Y_{t-1}\right] \;\leq\; \left(\frac{\beta_i}{n}\right)^2.$$

Prove that this implies

$$\mathbb{P}\left[\sum_{t=1}^{n} Y_t \geq \beta_{i+1}\right] \;\leq\; \mathbb{P}\left[\mathsf{Bin}\left(n, \left(\frac{\beta_i}{n}\right)^2\right) \geq \beta_{i+1}\right],$$

where $\mathsf{Bin}\left(n, p\right)$ denotes a binomial random variable with n independent trials and success probability p for each trial. Using Chernoff bounds, we get

$$\mathbb{P}\left[\mathsf{Bin}\left(n, \left(\frac{\beta_i}{n}\right)^2\right) \geq \beta_{i+1}\right] \;=\; \mathbb{P}\left[\mathsf{Bin}\left(n, \left(\frac{\beta_i}{n}\right)^2\right) \geq en \cdot \left(\frac{\beta_i}{n}\right)^2\right] \;\leq\; e^{-n \cdot (\beta_i/n)^2} \;\leq\; \frac{1}{n^2}$$

when $\beta^2 \geq 2n \ln n$. Thus,

$$\mathbb{P}\left[\neg E_{i+1} \mid E_i\right] \;=\; \frac{\mathbb{P}\left[\sim E_{i+1} \wedge E_i\right]}{\mathbb{P}\left[E_i\right]} \;\leq\; \frac{\mathbb{P}\left[\mathsf{Bin}\left(n, \left(\frac{\beta_i}{n}\right)^2\right) \geq en\left(\frac{\beta_i}{n}\right)^2\right]}{\mathbb{P}\left[E_i\right]} \;\leq\; \frac{1}{n^2} \cdot \frac{1}{\mathbb{P}\left[E_i\right]}$$

when $\beta^2 \geq 2n \ln n$. ∎

We can then use induction to show that for each i as above, the probability of the event E_i not happening is very low.

Claim 1.2 *For all i such that $\beta_i^2 \geq 2n \ln n$, we have*

$$\mathbb{P}\left[\neg E_{i+1}\right] \;\leq\; \frac{i+1}{n^2}.$$

Proof: We prove the claim by induction on i. We know from the definition of β_6 that $\mathbb{P}[\neg E_6] = 0$. Also, from the previous claim, we have that for any i as above,

$$\begin{aligned}\mathbb{P}[\neg E_{i+1}] &= \mathbb{P}[E_i] \cdot \mathbb{P}[\neg E_{i+1}|E_i] \;+\; \mathbb{P}[\neg E_i] \cdot \mathbb{P}[\neg E_{i+1}|\neg E_i] \\ &\leq \mathbb{P}[E_i] \cdot \frac{1}{n^2} \cdot \frac{1}{\mathbb{P}[E_i]} \;+\; \frac{i}{n^2} \\ &\leq \frac{i+1}{n^2}.\end{aligned}$$

■

We will need a slightly different analysis when $\beta_i^2 < 2n \ln n$. Let i_0 be the minimum i such that $\beta_i^2 < 2n \ln n$. Because $\beta_{i_0-1}^2 \geq 2n \ln n$, we have by the previous claim that $B_{i_0} \leq \beta_{i_0}$ with high probability. The probability that B_{i_0+1} is large can be bounded as before using

$$\begin{aligned}\mathbb{P}[(B_{i_0+1} \geq k) \wedge E_{i_0}] &\leq \mathbb{P}\left[\mathsf{Bin}\left(n, \left(\frac{B_{i_0}}{n}\right)^2\right) \geq k\right] \\ &\leq \mathbb{P}\left[\mathsf{Bin}\left(n, \left(\frac{\beta_{i_0}}{n}\right)^2\right) \geq k\right] \\ &\leq \mathbb{P}\left[\mathsf{Bin}\left(n, \left(\frac{2n \ln n}{n}\right)^2\right) \geq k\right],\end{aligned}$$

where we use the fact that the probability of seeing a certain amount of heads increases as we increase the probability of heads. If we set $k = 6 \ln n$, then Chernoff bound gives

$$\mathbb{P}[(B_{i_0+1} \geq 6 \ln n) \;\wedge\; E_{i_0}] \;\leq\; e^{-2 \ln n} \;=\; \frac{1}{n^2},$$

which implies as before

$$\mathbb{P}[(B_{i_0+1} \geq 6 \ln n)] \;\leq\; \mathbb{P}[(B_{i_0+1} \geq 6 \ln n) \;\wedge\; E_{i_0}] + \mathbb{P}[\neg E_{i_0}] \;\leq\; \frac{i_0+1}{n^2}.$$

We further look at whether there even exists a bin with load more than $i_0 + 2$, and we see that

$$\mathbb{P}[B_{i_0+2} \geq 1] \;=\; \underbrace{\mathbb{P}[B_{i_0+2} \geq 1 | B_{i_0+1} > k]}_{\leq 1} \cdot \underbrace{\mathbb{P}[B_{i_0+1} > k]}_{\leq \frac{i_0+1}{n^2}} \;+\; \mathbb{P}[B_{i_0+2} \geq 1 | B_{i_0+1} \leq k] \cdot \underbrace{\mathbb{P}[B_{i_0+1} \leq k]}_{\leq 1}.$$

Because B_{i_0+1} is small enough, it suffices to bound the only term left in the above equation with Markov's inequality,

$$\mathbb{P}[B_{i_0+2} \geq 1 | B_{i_0+1} \leq k] \;\leq\; \mathbb{E}[B_{i_0+2} | B_{i_0+1} \leq k] \;\leq\; \mathbb{E}\left[\mathsf{Bin}\left(n, \left(\frac{k}{n}\right)^2\right)\right] \;\leq\; \frac{k^2}{n}.$$

Recalling the expression for β_i

$$\beta_i = \frac{n}{2^{2^{i-6}} e},$$

we have

$$i_0 = \frac{\ln \ln n}{\ln 2} + O(1).$$

This completes the proof that if we choose two bins at random instead of one, we reduce the number of high-load bins from $O(\ln n)$ to $O(\ln \ln n)$ with high probability.

2 Martingales

We now relax the independence assumption we used in proving Chernoff-Hoeffding bounds. Martingale sequences capture the notion of somewhat limited independence which is still sufficient to prove similar concentration bounds. We first restate the a special case of Chernoff bounds slightly differently.

Let $X_1, \ldots, X_n$ be a sequence of independent random variables such that each X_i equals 1 with probability 1/2 and −1 with probability 1/2. Let

$$Z_i \;=\; X_1 + \cdots + X_i\,.$$

We take $Z_0 = 0$ and notice that Chernoff bounds imply that the difference $|Z_n - Z_0|$ is small with high probability. Note that the above sequence satisfies the property that

$$\mathbb{E}\left[Z_i \;\mid\; X_1, \ldots, X_{i-1}\right] \;=\; \mathbb{E}\left[Z_{i-1}\right]\,,$$

which turns out to be sufficient to prove the required concentration bounds. The sequence of random varaibles $\{Z_i\}_{i=1}^n$ is known as a *Martingale sequence*.

Definition 2.1 *Let $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_n$ be an inceasing sequence of σ-algebras, known as a* filter*, on a finite space Ω. A sequence of random variables $\{Z_i\}_{i=1}^n$ is known as* Martingale sequence *with respect to the above filter if for all $i \in [n]$, Z_i is measurable in the σ-algebra $\mathcal{F}_i$ and*

$$\mathbb{E}\left[Z_i \;\mid\; \mathcal{F}_{i-1}\right] \;=\; Z_{i-1}\,.$$

The sequence $Y_i = Z_i - Z_{i-1}$ is known as a martingale difference sequence*, and satisfies that*

$$\mathbb{E}\left[Y_i \;\mid\; \mathcal{F}_{i-1}\right] \;=\; 0\,.$$

Example 2.2 (Doob Martingale) *Let $A, X_1, \ldots, X_n$ be random variables on the same finite space Ω. Then check that*

$$Z_i \;=\; \mathbb{E}\left[A \;\mid\; X_1, \ldots, X_i\right]\,,$$

forms a martingale sequence. A case of particular interest is the one where $A = f(X_1, \ldots, X_n)$ is a function of the random variables $X_1, \ldots, X_n$.

We will prove a concentration inequality for such sequences, known as Azuma's inequality, in the next lecture.

References

[ABKU94] Yossi Azar, Andrei Z Broder, Anna R Karlin, and Eli Upfal, *Balanced allocations*, Proceedings of the twenty-sixth annual ACM symposium on Theory of computing, ACM, 1994, pp. 593–602.

[ABKU99] ———, *Balanced allocations*, SIAM journal on computing **29** (1999), no. 1, 180–200.

[Mit96] Michael David Mitzenmacher, *The power of two random choices in randomized load balancing*, Ph.D. thesis, PhD thesis, Graduate Division of the University of California at Berkley, 1996.