Proof of Azuma-Hoeffding Inequality

We upper bound $E[e^{\alpha(X_t - X_0)}]$.

Define $Y_i = X_i - X_{i-1}$, $i = 1, \ldots, t$.

Then, $X_t - X_0 = \sum_{i=1}^{t} Y_i$.

$|Y_i| \leq c$, and since $X_0, X_1, \ldots$ is a martingale,

$$E[Y_i | X_0, X_1, \ldots, X_{i-1}] = E[X_i - X_{i-1} | X_0, X_1, \ldots, X_{i-1}]$$

$$= E[X_i | X_0, X_1, \ldots, X_{i-1}] - X_{i-1} = 0.$$

We consider $E[e^{\alpha Y_i} | X_0, X_1, \ldots, X_{i-1}]$.

$Y_i = -c_i \frac{1 - Y_i/c_i}{2} + c_i \frac{1 + Y_i}{2}$

Since $e^{\alpha Y_i}$ is convex for all $\alpha > 0$,

$e^{\alpha Y_i} \leq e^{-\alpha c_i} \frac{1 - Y_i/c_i}{2} + e^{\alpha c_i} \frac{1 + Y_i/c_i}{2}$

$= \frac{e^{\alpha c_i} + e^{-\alpha c_i}}{2} + \frac{Y_i}{2c_i}(e^{\alpha c_i} - e^{-\alpha c_i})$.

Hence,

$$E[e^{\alpha Y_i}|X_0, X_1, \ldots, X_{i-1}]$$

$$\leq E[\frac{e^{\alpha c_i} + e^{-\alpha c_i}}{2} + \frac{Y_i}{2c_i}(e^{\alpha c_i} - e^{-\alpha c_i})|X_0, X_1, \ldots, X_{i-1}]$$
$$= \frac{e^{\alpha c_i} + e^{-\alpha c_i}}{2}$$
$$\leq e^{(\alpha c_i)^2/2}.$$

$$E[e^{\alpha(X_t - X_0)}] = E[\Pi_{i=1}^t e^{\alpha Y_i}]$$

$$E[E[\Pi_{i=1}^t e^{\alpha Y_i}|X_0, X_1, \ldots, X_{t-1}]]$$

$$= E[\Pi_{i=1}^{t-1} e^{\alpha Y_i} E[e^{\alpha Y_t}|X_0, X_1, \ldots, X_{t-1}]]$$

$$\leq E[\Pi_{i=1}^{t-1} e^{\alpha Y_i}] e^{(\alpha c_t)^2/2}$$
$$\leq e^{\alpha^2 \sum_{k=1}^t c_k^2/2}.$$

Thus,

$$\Pr(X_t - X_0 \geq \lambda) = \Pr(e^{\alpha(X_t - X_0)} \geq e^{\alpha\lambda})$$

$$\leq \frac{E[e^{\alpha(X_t - X_0)}]}{e^{\alpha\lambda}}$$

$$\leq e^{\alpha^2 \sum_{k=1}^{t} c_k^2/2 - \alpha\lambda}$$

$$\leq e^{-\frac{\lambda^2}{2\sum_{k=1}^{t} c_k^2}}$$

by choosing $\alpha = \frac{\lambda}{\sum_{k=1}^{t} c_k^2}$.

A similar argument gives the bound for $\Pr(X_t - X_0 \leq -\lambda)$.

Lipschitz condition

Let $f : D_1 \times \ldots \times D_n \rightarrow R$ be a real-valued function with n arguments from possibly distinct domains.

The function f is said to satisfy the **Lipschitz condition** with bound c, if for any $x_1 \in D_1, \ldots, x_n \in D_n$, any $i \in \{1, \ldots, n\}$, and any $y_i \in D_i$,

$$|f(x_1, \ldots x_{i-1}, x_i, x_{i+1}, \ldots, x_n) -$$

$$f(x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_n)| \leq c$$

That is, changing the value of any single coordinate can change the function value by at most c.

A General Framework for Applying Azuma's Inequality

Theorem 1. *Suppose we have a sequence of random variables* $X_1, \ldots, X_n$ *and a function* $f(\hat{X}) = f(X_1, \ldots, X_n)$ *which satisfies Lipschitz condition.*

Consider the Doob martingale:

$Z_0 = E[f(X_1, \ldots, X_n)]$ *and*

$Z_k = E[f(X_1, \ldots, X_n) | X_1, \ldots, X_k]$, $1 \leq k \leq n$.

If $X_1, \ldots, X_n$ *are independent random variables, then* $|Z_k - Z_{k-1}| \leq c$.

We can thus use Azuma to get a concentration bound for f.

Proof

We will show for discrete random variables that only take a finite number of values.

Let S_k stand for $X_1, X_2, \ldots, X_k$.

And $f_k(\hat{X}, x) = f(X_1, \ldots, X_{k-1}, x, X_{k+1}, \ldots, X_n)$.

and $f_k(\hat{z}, x) = f(z_1, \ldots, z_{k-1}, x, z_{k+1}, \ldots, z_n)$.

$Z_k - Z_{k-1} = E[f(\hat{X})|S_k] - E[f(\hat{X})|S_{k-1}]$

$Z_k - Z_{k-1} \leq max_x E[f(\hat{X}|S_{k-1}, X_k = x] - E[f(\hat{X})|S_{k-1}]$

$Z_k - Z_{k-1} \geq min_y E[f(\hat{X})|S_{k-1}, X_k = y] - E[f(\hat{X})|S_{k-1}]$

We bound:

$max_x E[f(\hat{X})|S_{k-1}, X_k = x] - min_y E[f(\hat{X}|S_{k-1}, X_k = y]$

$= max_{x,y}(E[f(\hat{X})|S_{k-1}, X_k = x] - E[f(\hat{X}|S_{k-1}, X_k = y])$

$= max_{x,y}(E[f_k(\hat{X}, x) - f_k(\hat{X}, y)|S_{k-1}])$

For any values $x, y, z_1, \ldots, z_{k-1}$,

$E[f_k(\hat{X}, x) - f_k(\hat{X}, y)|X_1 = z_1, \ldots, X_{k-1} = z_{k-1}]$

$= \sum_{z_{k+1},\ldots,z_n} \Pr((X_{k+1} = z_{k+1}) \cap \ldots \cap (X_n = z_n)).(f_k(\hat{z}, x) - f_k(\hat{z}, y)).$

$\leq c$.

Hence $E[f_k(\hat{X}, x) - f_k(\hat{X}, y)|S_{k-1}] \leq c$.

Balls and Bins

Suppose m balls are thrown independently and uniformly into n bins.

Let Z be the number of empty bins.

$\mu = E[Z] = n(1 - \frac{1}{n})^m \sim ne^{m/n}$.

We want to show a concentration result for Z.

Let X_t represent the bin into which the tth ball falls.

The sequence $Z_t = E[Z|X_1, \ldots, X_t]$ is a Doob martigale.

$Z = F(X_1, \ldots, X_n)$ satisfies the Lipschitz condition.

Hence, $|Z_t - Z_{t-1}| \leq 1$.

Thus, $\Pr(|Z - E[Z]| \geq \lambda) \leq 2e^{-\lambda^2/2m}$.