An Inequality on Guessing and Its Application to Sequential Decoding

ERDAL ARIKAN

Electrical Engineering Department, Bilkent University, 06533 Ankara, Turkey

Abstract — **Let (X,Y) be a pair of discrete random variables with X taking values from a finite set. Suppose the value of X is to be determined, given the value of Y, by asking questions of the form 'Is X equal to x?' until the answer is 'Yes.' Let $G(x|y)$ denote the number of guesses in any such guessing scheme when $X=x$, $Y=y$. The main result is a tight lower bound on nonnegative moments of $G(X|Y)$. As an application, lower bounds are given on the moments of computation in sequential decoding. In particular, a simple derivation of the cutoff rate bound for single-user channels is obtained, and the previously unknown cutoff rate region of multi-access channels is determined.**

I. The Inequality

Theorem 1 *For arbitrary guessing functions $G(X)$ and $G(X|Y)$, and any $\rho \geq 0$,*

$$E[G(X)^\rho] \geq (1+\ln M)^{-\rho}[\sum_{x\in\mathcal{X}} P_X(x)^{\frac{1}{1+\rho}}]^{1+\rho} \qquad (1)$$

and

$$E[G(X|Y)^\rho] \geq (1+\ln M)^{-\rho}\sum_{y\in\mathcal{Y}}[\sum_{x\in\mathcal{X}} P_{X,Y}(x,y)^{\frac{1}{1+\rho}}]^{1+\rho} \quad (2)$$

where $P_{X,Y}$, P_X are the probability distributions of (X,Y) and X, respectively, the summations are over all possible values of X, Y, and M is the number of possible values of X.

This result is a simple consequence of the following variant of Hölder's inequality.

Lemma 1 *Let a_i, p_i be nonnegative numbers indexed over a finite set $1 \leq i \leq M$. For any $0 < \lambda < 1$,*

$$\sum_{i=1}^{M} a_i p_i \geq \left[\sum_{i=1}^{M} a_i^{\frac{-\lambda}{1-\lambda}}\right]^{\frac{1-\lambda}{-\lambda}} \left[\sum_{i=1}^{M} p_i^{\lambda}\right]^{\frac{1}{\lambda}}$$

Proof. Put $A_i = a_i^{-\lambda}$, $B_i = a_i^{\lambda}p_i^{\lambda}$, in Hölder's inequality $\sum_i A_iB_i \leq \left(\sum_i A_i^{\frac{1}{1-\lambda}}\right)^{1-\lambda}\left(\sum_i B_i^{\frac{1}{\lambda}}\right)^{\lambda}$.

Proof of Theorem. Inequality (1) is obtained by taking $a_i = i^\rho$, $p_i = \Pr(G(X)=i)$, $\lambda = 1/(1+\rho)$ in the lemma, and noting that $\sum_{i=1}^{M} 1/i \leq (1+\ln M)$. Inequality (2) follows readily:

$$\begin{aligned} E[G(X|Y)^\rho] &= \sum_y P_Y(y)E[G(X|Y=y)^\rho] \\ &\geq \sum_y P_Y(y)(1+\ln M)^{-\rho}[\sum_x P_{X|Y}(x|y)^{\frac{1}{1+\rho}}]^{1+\rho} \\ &= (1+\ln M)^{-\rho}\sum_y[\sum_x P_{X,Y}(x,y)^{\frac{1}{1+\rho}}]^{1+\rho} \end{aligned}$$

II. Application to Sequential Decoding

To relate sequential decoding to guessing, let $\mathcal{X}$ denote the set of nodes in a tree code at some level N channel symbols into the tree from the tree origin. Let X be a random variable uniformly distributed on $\mathcal{X}$, indicating the node in $\mathcal{X}$ which lies on the transmitted path. Let Y denote the received channel output sequence when X is transmitted. Let $G(x|y)$ denote the rank order in which node $x \in \mathcal{X}$ is hypothesized (for the first time) by a sequential decoder when $X=x$ and $Y=y$. Moments of $G(X|Y)$ serve as measures of complexity for sequential decoding.

Let M be the size of $\mathcal{X}$, and $R = (1/N)\ln M$ denote the code rate. By Theorem 1 and the fact that $P_X(x) = 1/M$ for $x \in \mathcal{X}$, for $\rho > 0$,

$$E[G(X|Y)^\rho] \geq (1+NR)^{-\rho}\exp[\rho NR - E_0(\rho, P_X)]$$

where

$$E_0(\rho, P_X) = -\ln\sum_y[\sum_x P_X(x)P_{Y|X}(y|x)^{\frac{1}{1+\rho}}]^{1+\rho}.$$

Gallager [1, p. 149] shows that for discrete memoryless channels

$$E_0(\rho, P_X) \leq NE_0(\rho)$$

where $E_0(\rho)$ equals the maximum of $E_0(\rho, Q)$ over all single-letter distributions Q on the channel input alphabet. Thus, at rates $R > E_0(\rho)/\rho$, the ρth moment of computation performed at level N of the tree code must go to infinity exponentially as N is increased. The infimum of all real numbers R' such that, at rates $R > R'$, $E[G(X|Y)^\rho]$ must go to infinity as N is increased is called the cutoff rate (for the ρth moment) and denoted by $R_{cutoff}(\rho)$. We have thus obtained the following bound.

Theorem 2 *For any discrete memoryless channel with a finite input alphabet,*

$$R_{cutoff}(\rho) \leq E_0(\rho)/\rho, \quad \rho > 0. \qquad (3)$$

This result was proved earlier (for $\rho = 1$ only) in [2]; the present proof is much simpler. Moreover, the above method extends to the case of multiaccess channels, yielding their previously unknown cutoff rate region [3].

Acknowledgements

I am indebted to J.L. Massey and M. Burnashev for discussions on this problem.

References

[1] R. G. Gallager, *Information Theory and Reliable Communication.* New York: Wiley, 1968.

[2] E. Arikan, 'An upper bound on the cutoff rate of sequential decoding,' *IEEE Trans. Inform. Theory,* vol. IT-34, pp. 55-63, Jan. 1988.

[3] E. Arikan, 'An inequality on guessing and its application to sequential decoding,' submitted to *IEEE Trans. Inform. Theory,* Nov. 1994.